Proceedings of the Educational Research Conference 2001
Special Issue: Volume 2 Number 4
November 2001

Published by

Shannon Research Press
Adelaide, South Australia
ISSN 1443-1475
http://iej.cjb.net

International Education Journal

Volume 2, Number 4, Special Issue, November 2001

Proceedings of the Educational Research Conference 2001

Welcome to the International Education Journal. This is a broadly based journal encompassing research and review articles. Education is interpreted in a wide manner and includes human development, learning, school education, formal and informal education, tertiary and vocational education, industry training and lifelong learning.

Online - While there will be a hard copy version of the journal it is intended that it will be accessed in almost all instances via the web.

Peer Review - All articles appearing in this journal have been subjected to a peer-review process.

Published by:
Shannon Research Press,
13 Endeavour Court,
North Haven, SA 5018,
Australia

ISSN 1443-1475

Designed by Katherine Dix

Printed in Adelaide
March 2001

CONTENTS

FREE Online Subscription
Subscription is available as a free service in online format only. Either complete the subscription form at: http://iej.cjb.net or email the onlineeditor@iej.cjb.net. The reader will be notified of the latest issue and contents.

Publication Frequency

It is intended that the journal be published three times per year. However given the electronic characteristics of online publishing it is not necessary to wait until a full journal is assembled prior to publishing. Articles once accepted in appropriate form may be published without delay. It should be noted however that in electronic format pagination may vary according to the parameters of the downloading computer. Alternatively the journal article can be downloaded in PDF format, in which case the original pagination will be preserved.

CONTENTS

Aims of IEJ

The aim of the International Education Journal is to publish articles that possess one or more of the following characteristics or qualities.

1. The presentation of substantive findings which are of importance for policy and practice beyond the country in which the research was conducted.
2. The integration of education with academic disciplines such as anthropology, demography, economics, history, law, linguistics, philosophy, political science, psychology and sociology, or examines educational issues from the perspective of the disciplines or investigates issues at the interface between education and one or more of these disciplines.
3. The examination of educational issues from a cross-cultural or indigenous people's perspective.
4. The evaluation of educational policy or programs or the use of information technology of cross-national interest and significance.
5. The employment of advanced research methods and measurement procedures which are clearly explained.
6. The presentation of empirically or analytically based investigations of theory, models or conceptual frame works in the field of education.
7. The syntheses of research findings from comparative and cross-national studies in education.

Invitation

Contributors - Authors are invited tosubmit material to this journal. As a general guide articles should be not more than 5000 words. The journal may publish longer works such as theses or disserations as occassional mongraphs. In every instance material must be received in publisher ready format. Full details of publication style and other requirements are set out under the heading Author Information.

Publishers - If you wish to have books or educational software reviewed and published in the International Education Journal please contact the editors.

Reviewers - If you are interested in contributing to the journal by writing a review article (500-1000 words) please contact the authors indicating your areas of interest. We look forward to hearing from you.

CONTENTS

Author Information

IEJ welcomes practical and research manuscripts that focus on educational issues, provide a clear purpose and depth of discussion, and are resented in a straight-forward style accessible to an international audience.

The initial submission:

- should be prepared using the downloadable Microsoft WORD97 template document (iejbase.doc), complete with layout styles, images and tables correctly positioned in the text
- images that are imported or inserted into the text from an external source as figures should be included as separate graphic files (either GIF or JPG format)
- should be forwarded electronically, either by:

  e-mail to: onlineeditor@iej.cjb.net

  or,

  on a 1.4MB 3.5 inch floppy diskette to: IEJ Online Editor,

  39 Dorrien Avenue, Woodcroft,

  South Australia, 5162

The editors reserve the right to return poorly edited or improperly formatted manuscripts. Hard-copy only submissions cannot be accepted or returned.

The manuscript should:

- show the article's title, authors' names, affiliation, and email addresses (these will be included on the online page), and postal address and telephone numbers (these will be removed from the final online page)
- abstract should not exeed 150 words
- paper should not exceed 5,000 words (20,000 for occasional papers)
- include an abstract of about150 words
- include 5 keywords describing article
- be in final form ready for reviewing
- conform to the APA Reference System

General Principles for layout:

- Papers should be prepared using the template *iejbase.doc* with Microsoft WORD on Mac or PC
- Authors with WORD 98 or later and a basic understanding of web page construction should also save their paper in web format (**.html*) incorporating any hypertext features.
- Use end-notes, with one set of numbers running through the whole article.
- Use italics (not underline) for emphasis.
- Do not create additional headers and footers in the template document *iejbase.doc*, just modify appropriately.
- Do not use page numbers in text (web pages do not follow conventional page numbering systems).

CONTENTS

If you have any questions regarding layout and presentation, please contact the IEJ Online Editor by email on: onlineeditor@iej.cjb.net

FUIIE

The publishers have arranged for the Flinders University Institute of International Education to produce and host an electronic mirror site containing all material in the Journal, accessed free of charge. There are no restrictions on downloading or printing single copies. The electronic version also provides readers with a search tool.

The Educational Research Conference 2001

IEJ is proud to support the Proceedings of the Educational Research Conference 2001, held annually at the Flinders University School of Education.

The Flinders University Institute of International Education and the School of Education co-sponsor the conference with the Graduate School of Education, University of Adelaide.

The one-day conference brings together staff and postgraduate students from the three South Australian Universities to share research findings and hear guest speakers.

Held on the 27th November 2001, the guest speakers were Professor Jan-Eric Gustafsson from Goteborg University Sweden, presenting his work on schooling and intelligence, and Dr Robyn Young from Department of Psychology Flinders University, speaking about Autism and Savant Syndrome.

Measuring cognitive and non-cognitive systems of reasoning: some preliminary findings

Carol R. Aldous
Flinders University *carol.aldous@flinders.edu.au*

This paper reports on preliminary findings involving a pilot project for a doctoral thesis by research, which seeks to examine the issues of creativity in problem solving and of how such creativity may be fostered in children under instruction. In particular, the design, trial and statistical appraisal of a new self-report instrument (viz: Systems of Reasoning Questionnaire SRQ) that was formulated to assist in the identification, description and measurement of some cognitive and non-cognitive forms of reasoning is described. The SRQ comprises five scales, which tap into Rule-based and Associative forms of reasoning as characterized by Sloman, 1996. The pilot project involved sampling upper primary and lower secondary school students solving novel mathematics problems within the Mathematics Challenge for Young Australians 2000.

Cognitive, Non-cognitive, Rule-based, Associative, Creativity

INTRODUCTION

Recent critique of educational research related to constructivist thinking, particularly with regard to mathematics education, has pointed to gaps in understanding of the part cognitive and non-cognitive systems of reasoning play in solving problems (Davis 1996, Davis et al 1996, Taylor P.1996, Begg 1999, Gunn 1999). Indeed concern has been expressed that constructivist approaches to mathematics teaching and learning focus largely on cognitive forms of reasoning and take little account of other kinds of knowing such as those of how feeling and emotion are constructed in the learning and/or the problem solving process, of how non-conscious or unformulated knowing may contribute, or of how things may be known intuitively.

Given that for the past twenty five years at least, cognitive psychology (unlike its counterpart of social psychology) has largely ignored affect (Zajonic, 1980, Anderson, 1999) and that many current models of constructivist teaching and learning have their early beginnings in cognitive and experimental psychology (Good, Wandersee & St Julien 1993; Phillips 2000), the concerns mentioned above would appear to be entirely consistent and largely inevitable. Since teaching and learning must of necessity, centre on the total individual, educational practices which focus solely on cognitive aspects to the exclusion of non-cognitive ones, are in the long term, likely to be found wanting (Begg 1999). In fact these concerns have been finding expression in literature for some considerable time. As far back as 1976 for instance, Miller and Johnson-Laird in their notable work on language and perception in which "fearfully cognitive and dispassionate"(p.111) information–processing systems of perception (*Perceive*), memory (*Remember*) and intention (*Intend*) were presented, had this to say about feeling (*Feel*):

> *Feel* is an indispensable predicate for any complete psychology and that it probably lies much closer than *Perceive, Remember,* and *Intend* to the basic sources of energy that keep the whole system running (p.112) (Miller and Johnson-Laird italics)

In 1988, Torrence and Rockstein elaborating on future directions in research and teaching wrote:

> As a first step in these directions, teachers and teacher educators must work together to develop whole-brain units of learning that combine objectives from the cognitive, affective and intuitive domains (p.289)

In 1999, mathematics educator, Begg writing about non-cognitive knowing in the context of research on learning, teaching and curriculum stated:

> Whether we consider emotions as unformulated knowledge, as personal constructions, or as actions, there seems to be a need to consider them not apart from other constructions and actions but as part of our being which is intimately interrelated with our knowing. (p.70)

Perfinking and Creativity

Highlighting the idea that individuals may, perceive, feel and think all at once, Bruner (1986) cites David Krech as coining the term "perfink" and goes on to add that action, is an important component of "perfinking" (Bruner 1990). Indeed, perceiving, feeling, thinking and acting are all likely to be present during learning (Davis 1996). Thus in the endeavour to assist children solve unseen mathematical problems as is the case in novel problem situations, the capacity to draw upon "perfinking" not merely cognitive thinking, may well increase the potential for the enactment of a successful and creative solution.

Certainly teaching children to tap into both their cognitive and non-cognitive forms of reasoning holds significance in the field of creativity research (Russ 1993; Sternberg and O'Hara 1999) where some operational definitions of creativity embrace the notion of both subjective and objective knowing (Taylor 1989). Research in this field has shown that when students are taught and assessed in a way that values both their subjective and objective creative abilities their academic performance improves (Sternberg et al 1996; Sternberg et al 1998). In seeking to make explicit the nature of the creative process Koberg and Bagnall (1976) for example described creativity as:

> … both the art and the science of thinking and behaving with both subjectivity and objectivity. It is a combination of feeling and knowing: of alternating back and forth between what we sense and what we already know. Becoming more creative involves becoming awake to both; discovering a state of wholeness which differs from a primarily objective or subjective person which typifies our society. (p.8)

While much has been written about the importance of creativity to novel problem solving (Sternberg 1999) and terms such as "perfinking" point to gaps in approaches to teaching and learning (Davis 1996), little has been done to address these issues in education generally and in mathematics education in particular. In point of fact, apart from social psychological consideration such as awareness that feelings of anxiety may affect problem solving ability, non-cognitive forms of reasoning have generally gone un-noticed in the teaching of students to solve problems creatively. This is partly because the identification, description and measurement of non-cognitive forms of reasoning in the schooling context have largely been illusive and partly because their interaction with cognitive forms of reasoning remains uncertain. While the creativity literature has many examples of individuals drawing upon the elements of imagination, intuition and feeling in creating eminent works, including mathematicians (Parr, 1974; Morris, 1993; Simonton, 1999; Feist, 1999; Policastro & Gardner 1999) little if any research has been done to uncover these forms of reasoning amongst students under mathematics instruction.

Hence it is the purpose of this research to determine whether cognitive and non-cognitive forms of reasoning can be identified and described in the novel mathematics problem-solving context and whether they can be measured through a self-reporting instrument designed to tap into two kinds of reasoning. The remainder of this paper outlines the methods used, the reasons for selecting these methods and the resulting findings.

METHODOLOGY

In 1993 Biggs presented a model of teaching and learning, (termed the '3P' model of teaching and learning) which provides a method for separating and assessing various elements in a teaching and learning situation. Central to this model is the notion that by measuring a student's approach to a given learning activity it is also possible to tap into other elements of the teaching and learning system such as its presage, process and product (i.e. the three Ps) (Biggs 2001). The presage component may be determined by student factors and by the teaching context, the process component by the nature of the learning activities and the product component, by the learning outcomes (Biggs 2001).

In similar vein, it was conjectured that if it is possible to measure a student's reasoning approaches to a novel problem solving activity, it may also be possible not only to tap into the process(es) (both cognitive and non-cognitive) associated with creative problem solving, but also, to access in the long term other elements of the teaching and learning system which may influence student creativity in a mathematical problem solving context. One analgous model suggests creativity may be determined by the intersecting set of the person, process, product and environment (Aldous 1999, 2000). As Shuell (1986) explained "It is important to remember that what the student does is more important than what the teacher does." (p.429)

Determining the Learning Activity

The research of these possibilities and conjectures clearly required a student learning and problem solving activity on a broad basis. Furthermore the learning activity also needed to be concomitant with a four stage model, of creative problem solving frequently depicted by psychologists, mathematicians and scientists alike (Parr 1974, Russ 1993). (This model was first described in the early writings of Wallas (1926) and Hadamard (1945)). In brief the four stages of creative problem solving include: *preparation, incubation, illumination* and *elaboration/verification*. In addition the learning activity had the indispensable requirements that it be novel, if creativity were to be expressed (Runco & Sakamoto 1999) and that the activity should be undertaken by a sufficiently wide-ranging sample of students to be statistically both useful and fruitful (Keeves 1997).

Thus the student learning activity central to the framework for research was determined to be the Mathematics Challenge for Young Australians, a national competition of novel mathematics problem solving for upper primary and lower secondary students. The activity, an annual event, has the unique feature of allowing students time to reflect and to incubate on novel problems by providing a challenge time of any three weeks, during a specified six week period, in which to answer six novel problems. The time component of the Challenge was considered vital, since it was more likely to provide students with the opportunity to demonstrate cognitive and non-cognitive forms of reasoning. Moreover the notion of novelty in the problem-solving context was also critical. Much problem solving that takes place in mathematics classrooms is typically not novel but the Mathematics Challenge afforded the opportunity that the problems the students were to encounter were more likely to be new and different to them. Hence creative thinking and cognitive and non-cognitive ways of reasoning were more likely to be expressed. As Getzels (1975) stated "… the creativity of a solution depends on the creativity of the problem being solved."(p.84). Finally, because the Australian Mathematics Challenge was a national event access to a wide ranging sample from a broad cross section of the community was possible.

It was against the background of this national problem solving challenge event that a self-report instrument was designed in an attempt to measure two kinds of reasoning in creative problem solving.

Designing the Systems of Reasoning Questionnaire (SRQ)

In designing an instrument, which might be used to tap into cognitive and non-cognitive forms of reasoning a search of the relevant literature drew evidence from neurobiology, empirical psychology, popular psychology and philosophy. However for the confines of this paper, discussion will be restricted to evidence drawn predominantly from the fields of empirical and popular psychology.

Two Systems of Reasoning

Popular Psychology

Popular psychology has for some time now talked about two kinds of thinking or reasoning. The first kind has been popularly termed "computer thinking" and the other "free-flowing" or "transmitter thinking" (Carlson and Bailey 1997; Carlson 1998). Computer thinking has been characterised as effortful, linear, active, memory based and memory bound, conditional on emotions, predictable, and task detailed. Free-flowing thinking on the other hand has been characterised as passive, non-linear, selective of memory, employing deeper feelings, spontaneous, creative, inspired, visionary and big picture oriented.

Computer thinking has been popularly described as the kind of thinking which might be used by an individual in planning the steps needed to get from home to the airport in order to catch the aeroplane on time. It is thought to involve that part of the brain used in analysing, comparing, relating facts and making computations and is the kind of thinking generally taught in schools. Free-flowing thinking on the other hand has been popularly described as the kind of thinking that may occur when the individual does not know all the variables implicit to a problem, but having considered as many aspects as possible, sets aside the problem to incubate in the subconscious mind. Should the solution be required by a specified time, so long as the conscious mind has been primed then an answer should intuitively manifest itself in the time required (Carlson and Bailey 1997).

Hence it is from descriptions such as these that it may be possible to extrapolate that the popularly described computer thinking and free-flowing thinking are indicative to some degree of cognitive and non-cognitive forms of reasoning.

Empirical psychology

While the embracement of such popularly described forms of thinking should be treated with caution, the concept of two kinds of reasoning is well grounded in empirical psychology (Sloman 1996). Indeed the differentiation of two systems of reasoning have been cited as far back as Aristotle and can be found in the writings of James (1890/1950), Piaget (1926), Vygostsky (1934/1987) and Bruner (1986). While the configuration of the two systems may have shifted emphasis throughout history, with some psychologists arguing for one type of thinking in preference for another (eg. parallel processing through associative pathways versus symbol processing through serial analysis (Baer 1993)), others have argued for the dual existence of both systems of thinking (Neisser 1963; Johnson Laird 1983).

One advocate who conceives of the mind operating in both ways is Sloman (1996) and it is upon his empirical characterization of two forms of reasoning, that the Systems of Reasoning Questionnaire (SRQ) is based. (A summary of Sloman's characterisation of two systems of reasoning is provided in Table 1.)

Instead of names such as free-flowing and computer thinking, Sloman terms the dual systems of reasoning as the "associative system" and the "rule-based system". In comparing the associative system with the rule-based system he explains:

> Associative thought *feels* like it arises from a different cognitive mechanism than does deliberate, analytical reasoning. Sometimes conclusions simply appear at some level of awareness, as if the

> mind goes off, does some work, and then comes back with a result, and sometimes coming to conclusion requires doing the work oneself, making an effort to construct a chain of reasoning. (p.3, Sloman's italics)

Table 1. A table taken from page 7 of Sloman, S. A. (1996). The Empirical Case for Two Systems of reasoning. Psychological Bulletin, 119(1) 3-22 is shown.

TWO FORMS OF REASONING

Table 1
Characterization of Two Forms of Reasoning

Characteristic	Associative system	Rule-based system
Principles of operation	Similarity and contiguity	Symbol manipulation
Source of knowledge	Personal experience	Language, culture, and formal systems
Nature of representation		
Basic units	Concrete and generic concepts, images, stereotypes, and feature sets	Concrete, generic, and abstract concepts; abstracted features; compositional symbols
Relations	(a) Associations	(a) Causal, logical, and hierarchical
	(b) Soft constraints	(b) Hard constraints
Nature of processing	(a) Reproductive but capable of similarity-based generalization	(a) Productive and systematic
	(b) Overall feature computation and constraint satisfaction	(b) Abstraction of relevant features
	(c) Automatic	(c) Strategic
Illustrative cognitive functions	Intuition Fantasy Creativity Imagination Visual recognition Associative memory	Deliberation Explanation Formal analysis Verification Ascription of purpose Strategic memory

Using an example from arithmetic to illustrate the difference, Sloman continues:

> Given an arithmetic problem such as figuring out change at the cash register, sometimes the answer springs to mind associatively, and sometimes a person has to do mental arithmetic by analyzing the amounts involved and operating on the resultant components as taught in school. (p.3)

In more precise discussion based on empirical evidence, Sloman characterises associative reasoning as operating reflexively, through a kind of statistical description of the environment in which similarity between problem elements is used to draw inferences. Such inferences may be drawn from general knowledge in the form of images, templates and stereotypes. Rule-based reasoning on the other hand operates in a causal-mechanical manner and is interpreted through different kinds of logical and hierarchical structures.

Notably, Sloman identifies a number of cognitive functions which are illustrative of the expert capacity of each reasoning system. The descriptions of many of these functions were used in framing statements for the SRQ. Intuition, creativity, imagination, fantasy, visual recognition and associative memory are the cognitive functions, which typify associative reasoning. Deliberation, explanation, formal analysis, verification, ascription of purpose and strategic memory are the cognitive functions, which typify rule-based reasoning. (Refer items 19, 20, 27 and 29 to 37 for some examples). Furthermore the source of knowledge in associative reasoning is derived from personal experience while that of rule-based reasoning is derived from formal elements in the culture. Thus the characterization of a student's knowledge

source was also useful in formulating items for the SRQ and thereby the possible identification of the system of reasoning being used. (Items 21, 23 and 24 are cases in point.)

It should be pointed out that Sloman does not suggest in his findings that one style of reasoning exists in isolation from the other. Indeed they frequently overlap. Individuals may use both forms of reasoning in solving a problem even to the extent that conflicting or different answers may result. Drawing upon empirical data Sloman (1996) shows that individuals may simultaneously believe two contradictory answers to the same reasoning problem to demonstrate this point. Consequently identifying and classifying the two reasoning systems is not without its difficulties. Hence, to the degree in which individuals may use both forms of reasoning at the same time, some items in the SRQ may be characterized as drawing upon both rule-based and associative forms of reasoning, although in most instances there is a tendency to favour one form over another. (Items, 44, 45, 48 and 49 are examples of this.)

Sloman goes on to explain that each type of reasoning need not be exclusive to a given problem domain but that each mode may interact with the other contributing different computational resources to the problem at hand. Indeed because a given type of reasoning cannot be identified by a given problem domain, determining which system is responsible for a given response is not always possible. However, as a general rule the "contents of awareness"(p.6) Sloman (1996) states can be used as an indicator. Should an individual solve a problem with little or no awareness of the process used in the solution but be conscious only of the result then the response is likely to have been produced by the associative system. Thus a number of items were developed which sought to probe a student's content of awareness and thus give a likely indication of which system of reasoning is being used. (Items 15,16, 17 and 18 are cases in point).

A table documenting the forty-five statements developed for the Systems of Reasoning Questionnaire (SRQ) together with the source of the idea for each statement is shown in Table 2. In addition the form of reasoning, be it rule-based or associative, which each statement was intended to measure is also given. The majority of these statements may be mapped to one of the five categories (viz: *Principles of Operation, Source of knowledge, Nature of Representation, Nature of Processing and Illustrative Cognitive Function*) outlined by Sloman in Table 1. In addition the category "contents of awareness" was added for emphasis, although statements in this category may be mapped to *Nature of Processing*. Furthermore ten items were also formulated which related to the four stage model of creative problem solving elaborated by Russ (1993). It was hoped that some pattern could be found linking the stages of creative problem solving with the form of reasoning used.

Procedure

Pilot Sample

A sample of 114 students from across four metropolitan private schools ranging from year 7 to year 10 was used in this pilot study. Of this sample, 56 students were male and 58 female. Subjects in each school were selected on the basis of their entry into the 2000 Australian Mathematics Challenge and on their willingness to be involved in the study. Participation in the study required that the student complete the questionnaire twice, once for each of the two problems selected, one in *number* and one in *space*. Thus although the challenge comprised six novel problems, students involved in the study answered the SRQ on only two designated problems.

Table 2. A table documenting the origin of ideas contained within each statement and the form of reasoning viz: associative and/or rule-based it was anticipated to measure in SRQ.

Item	Statement	Source of Idea; Author	Forms
15	I got answer first, then thought out mathematical reasons for it being correct	Contents of awareness; Sloman 96	Associative
16	I got the answer by using mathematical steps right from the beginning	Contents of awareness; Sloman 96	Rule Based
17	I got the answer but then had trouble explaining how I arrived at it	Contents of awareness; Sloman 96	Associative
18	I just knew how to get answer without using mathematical reasoning first	Contents of awareness; Sloman 96	Associative
19	I used my imagination in solving this problem	Cognitive function; Sloman 96	Associative
20	I used my intuition (followed a gut feeling) solving this problem	Cognitive function; Sloman 96	Associative
21	I used things I learned in schools while solving this problem	Source of knowledge; Sloman 96	Rule Based
22	The “answer” or “How to get the answer” suddenly came into my head while I was working on this problem	Nature of processing; Sloman 96	Associative
23	I associated this problem with things I experienced outside of school	Source of knowledge; Sloman 96	Associative
24	I associated this problem with images/pictures/diagrams I have seen	Source of knowledge; Sloman 96	Associative
25	I saw spatial/visual patterns in my mind	Nature of representatn; Sloman 96	Associative
26	I recognized number patterns in solving this problem	Nature of representatn; Sloman 96	Rule Based
27	I felt that I related to /connected with the patterns in this problem	Cognitive function; Sloman 96	Associative
28	I had a sense of number size in this problem	Nature of representatn; Sloman 96	Associative
29	I organized my reasoning /thinking in a strategic way	Cognitive function; Sloman 96	Rule Based
30	I recalled specific maths facts in solving this problem	Cognitive function; Sloman 96	Rule Based
31	I used a sequence of logical steps in this problem	Cognitive function; Sloman 96	Rule Based
32	I used a strategy/procedure I learned in class to solve this problem	Cognitive function; Sloman 96	Rule Based
33	I used specific formulae in solving this problem	Cognitive function; Sloman 96	Rule Based
34	I tried to verify/check that my answer was correct	Cognitive function; Sloman 96	Rule Based
35	I let my mind go free, thinking of any possibility solving this problem	Cognitive function; Sloman 96	Associative
36	I drew upon all mental resources/parts of me to solve this problem	Cognitive function; Sloman 96	Associative
37	I tried to be inventive in solving this problem	Cognitive function; Sloman 96	Associative
38	I developed a feeling about correctness of my solution before I checked it	Nature of processing; Sloman 96	Associative
39	I thought of how to get the answer to this problem while I was doing something else (eg working another problem, riding a bike)	Nature of processing; Sloman 96	Associative
40	I thought of how to get answer to problem when I woke up in morning	Nature of processing; Sloman 96	Associative
41	I tried a lot of different ways until I found the right one	Cognitive function; Sloman 96	Rule Based
42	Thought mainly with specific and exact numbers	Nature of representatn; Sloman 96	Rule Based
43	Thought mainly with approximate and general numbers	Nature of representatn; Sloman 96	Associative
44	Thought with both exact and approximate numbers	Nature of representatn; Sloman 96	Ass. & Rb
45	Thought with pictures/diagram (ie spatially) as well as with words	Nature of representatn; Sloman 96	Ass. & Rb
46	Thought more with words/symbols than with pictures/diagrams	Nature of representatn; Sloman 96	Rule Based
47	Thought more with pictures/diagrams than with words/symbols	Nature of representatn; Sloman 96	Associative
48	Thought with pictures & words/symbols at the same time	Nature of representatn; Sloman 96	Ass. & Rb
49	Often alternated between pictures and words	Nature of representatn; Sloman 96	Ass. & Rb
50	I carefully read the problem more than once	Creative PS Preparation; Russ 93	Rule Based
51	I took time to understand what the problem was asking	Creative PS Preparation; Russ 93	Rule Based
52	I had several attempts at finding a solution before I gave up for a time and came back to it later	Creative PS Incubation; Russ 93	Associative
53	I played around and explored the problem for a while	Creative PS Preparation; Russ 93	Rule Based
54	I spent time reflecting on how to solve the problem	Creative PS Incubation; Russ 93	Rule Based
55	I followed a feeling/hunch about what to do	Creative PS Illumination; Russ 93	Associative
56	I set the problem aside for a time	Creative PS Incubation; Russ 93	Associative
57	Having set aside the problem for a time I found the solution suddenly popped into my mind	Creative PS Illumination; Russ 93	Associative
58	The solution occurred to me after I had thought consistently for a time	Creative PS Illumination; Russ 93	Rule Based
59	I checked that my solution was correct	Creative PS Verification; Russ93	Rule Based

Instructions to Participants

Because entrants in the Mathematics Challenge had a period of three weeks in which to answer six challenge problems, participants were instructed to complete the SRQ as soon as practicable after completion of the relevant problem in order to facilitate accuracy and aid their responding. With regard to these instructions the scale devised for each of the 45 SRQ statements on which participants were invited to respond, comprised three points. These were True, Not Sure and False. The selection of a trichotomous scale was done for simplicity in order to keep the amount of missing data to a minimum (Keeves 1966) while at the same time providing a central category for participants who genuinely wished to indicate uncertainty on the grounds that the extreme category did not apply in its entirety. In addition it was felt that having fewer choices would be less confusing for the younger students in the sample, alleviating the dilemma associated with Likert scales of five points or more, in which some individuals tend to answer in the extremes while others routinely avoid them (Kline 1993). Further should it be necessary to unpack the responses amongst participants who responded with the "not sure" category (viz: "not sure" because they did not understand the statement or "not sure" because they were genuinely uncertain in their answer) the Hyperbolic Cosine Model (HCM) for unfolding responses could be used. (Andrich and Luo 1993, Andrich 1996). In the subsequent analysis, a score of two was assigned for a response of true, one for uncertain and zero for false.

Analysis Overview

The data collected in this pilot study were subjected to principal-components factor analysis (i.e. exploratory factor analysis) using SPSS (Version 10.0) rather than the more powerful confirmatory factor analysis for several reasons. Firstly, since research into the identification and description of cognitive and non-cognitive forms of reasoning during creative problem solving remains largely unoperationalised and the instrument devised in this study for measuring two forms of reasoning is in its infancy, a procedure which allowed the data to determine the underlying factor model was to be preferred in the first instance, over one that dictated what the model should be. As Carroll (1983) explains:

> Generally, exploratory methods are to be preferred for establishing the probable factorial composition and structure of a set of variables; confirmatory methods can then be applied to establish or test their significance. (p. 15)

Thus use of the exploratory factor analysis procedure in this study, was seen as a tool for theory building (Bryant and Yarnold 1998), which may later be put to the test using confirmatory factor analysis. Secondly the size of the data sample (N= 114) meant that conditions for using confirmatory factor analysis were less than optimal (Gustafsson and Balke 1993). Confirmatory factor analysis will be employed in the follow up study of the SRQ using a sample of at least N=400.

Initial Analysis

The exploratory factor analysis of the SRQ proceeded in several distinct stages. First because the SRQ was answered on two separate problems one in *number* and one in *space* separate analyses were conducted for each problem to determine any similarities and or differences in the formation of factors and to assess whether the data could legitimately be combined to yield a larger number of observations. Following the general principle of a *subjects–to-variables* (STV) ratio of approximately five or greater (Bryant and Yarnold 1998) these initial analyses were conducted on 27 items out of a possible 45 from the SRQ. Each data set was subjected to a principal components extraction followed by a varimax rotation, yielding three distinct but common factors. These factors were recognizable in both *number* and *space* as the *Systematic* Factor, the *Free-flowing* Factor and the *Feeling* Factor. Although there are differences in the distribution and loading of some items, the factors were found by inspection to be sufficiently stable (i.e. similar and constant) across problem type for the data sets to be combined for

subsequent analyses. Furthermore since the focus of the study is on the processes of creative problem solving rather than the influence of problem type and that these processes are hypothesized to be generic across problem situations the final analyses were conducted on the combined space and number data and it is upon the combined data that the paper reports. The results of these initial rotated factor solutions are located in Appendices 1 and 2.

Final Analysis

Following the initial exploratory factor work, principal components factor analyses were now conducted on 45 items in the SRQ using the observations from the combined space and number data (N = 198, listwise deletion of data). These were followed by a varimax rotation to achieve simple structure, assisting the resolution of factors that would help build a model of creative problem solving involving cognitive and non-cognitive forms of reasoning. The results of the rotated factor pattern are given in Tables 3 and 4.

RESULTS AND DISCUSSION

Factor Structure

A five-factor structure accounting for 38 percent of the total variance was obtained. Although initial principal components analysis yielded 13 components whose eigenvalues were greater than one, accounting for 62 percent of the variance, five of the 13 components had eigenvalues greater than two. A scree test (Cattell 1966), performed in order to determine the number of factors, clearly indicated the presence of five distinct factors. A varimax rotation of five factors was undertaken and of the 45 items under consideration in the SRQ, 43 had loadings of 0.30 and above. Further a six-factor extraction carried out as a precautionary measure, yielded un-meaningful data.

With the exception of two items (viz: item 35 on factor 1 and item 46 on factor 3) all items within each scale are positive. However taking into consideration the composition of all other items in each respective scale, the negative loadings for these items is not inconsistent with the interpretation of each scale. Nonetheless consideration as to whether these items are ultimately included will be undertaken upon further trailing. The items, which did not load according to the 0.3 criteria, were item 27 (*I felt that I related to/connected with the patterns in this problem (0.27 on factor 5))* and item 43 (*Thought mainly with approximate and general numbers (0.26 on factor 4)).*

The reliability, for each factor (Kaiser and Caffrey 1965) is documented in Table 4, together with the percent variance explained by each factor, both before and after orthogonal rotation. Factors one, two and three display good reliability while factors four and five have satisfactory reliabilities.

Interpretation of Factors and the Formation of Scales

On the basis of the factorial structure arrived at, five scales corresponding to each of the five factors have been derived. These scales have been identified as:

- *Systematic* approach to reasoning
- *Strategic* approach to reasoning
- *Spatial/Verbal* approach to reasoning
- *Free-flowing* approach to reasoning
- *Feeling* approach to reasoning

Table 3. Principal Component Solution Rotated to Orthogonal Structure (loadings > 0.3)

Item	Statement	Fac1	Fac2	Fac3	Fac4	Fac5
A28	I had a sense of number size in this problem	0.57				
R29	I organized my reasoning /thinking in a strategic way	0.57				
R26	I recognized number patterns in solving this problem	0.56		-0.33		
R16	I got the answer by using mathematical steps right from the beginning	0.53				
R30	I recalled specific maths facts in solving this problem	0.53				
R31	I used a sequence of logical steps in this problem	0.52				
R33	I used specific formulae in solving this problem	0.51				
RA44	Thought with both exact and approximate numbers	0.47				
R21	I used things I learned in schools while solving this problem	0.46				
R42	Thought mainly with specific and exact numbers	0.45				
R32	I used a strategy/procedure I learned in class to solve this problem	0.38				
A35	I let my mind go free, thinking of any possibility solving this problem	-0.36				
R41	I tried a lot of different ways until I found the right one		0.62			
R51	I took time to understand what the problem was asking		0.61			
R53	I played around and explored the problem for a while		0.60			
R50	I carefully read the problem more than once		0.55			
A17	I got the answer but then had trouble explaining how I arrived at it		0.51			
A52	I had several attempts at finding a solution before I gave up for a time and came back to it later		0.50			-0.33
R58	The solution occurred to me after I had thought consistently for a time		0.47			
R54	I spent time reflecting on how to solve the problem		0.47			
AR 45	Thought with pictures/diagram (ie spatially) as well as with words			0.78		
A47	Thought more with pictures/diagrams than with words/symbols			0.77		
R46	Thought more with words/symbols than with pictures/diagrams			-0.61		
AR49	Often alternated between pictures and words			0.57	0.34	
A25	I saw spatial/visual patterns in my mind			0.56		
A24	I associated this problem with /images/pictures/diagrams I have seen			0.53		
AR48	Thought with pictures & words/symbols at the same time			0.33		
A23	I associated this problem with things I experienced outside of school				0.59	
A40	I thought of how to get answer to problem when I woke up in morning				0.59	
A57	Having set aside the problem for a time I found the solution suddenly popped into my mind				0.56	
A39	I thought of how to get the answer to this problem while I was doing something else (eg. working another problem, riding a bike)				0.56	
A36	I drew upon all mental resources/parts of me to solve this problem				0.38	
A56	I set the problem aside for a time		0.31		0.36	-0.33
R34	I tried to verify that my answer was correct		0.42			0.59
A55	I followed a feeling/hunch about what to do					0.59
R59	I checked that my solution was correct		0.33			0.59
A20	I used my intuition (followed a gut feeling) solving this problem					0.48
A19	I used my imagination in solving this problem	-0.34				0.47
A22	The “answer” or “How to get the answer” suddenly came into my head while I was working on this problem					0.45
A18	I just knew how to get the answer without using maths* reasoning first		-0.38			0.45
A38	I developed a feeling about correctness of my soln* before I checked it	0.35				0.41
AR15	I got answer first, then thought out maths* reasons for it being correct					0.38
A37	I tried to be inventive in solving this problem				0.31	0.36

Factor one, which corresponds to the scale for measuring *Systematic* approaches to reasoning, reflects the intent of a student who structures his thinking in a methodical and systematic way, who consciously and deliberately seeks to recall relevant information (e.g. number facts,

procedures) to solve a problem and who progresses his thinking in a linear, logical and effortful way.

Table 4. Summary information for each scale

Factor No.	Scale Name	Nos. of Items	Reliability Kaiser Caffrey	% Variance explained before rotation	% Variance explained after rotation
One	Systematic	12	0.75	11.09	8.34
Two	Strategic	8	0.70	8.39	7.68
Three	Spatial/Verbal	7	0.76	6.98	7.52
Four	Free-flowing	6	0.65	6.00	7.23
Five	Feeling	10	0.68	5.13	6.83

Systematic approaches to reasoning are largely the kinds of thinking processes, which have generally been emphasized in mathematics curricula as taught in schools (Lovitt 1999). While these kinds of thinking strategies are critical to the documentation of a solution and ultimate success, of themselves, they may yet be found to be insufficient in fruitfully solving a truly novel problem.

Factor two, which has been identified as the *Strategic* approach to reasoning, reflects the tactics of a student who employs elements of the four stage model of creative problem solving to facilitate their thinking including conscious and deliberate preparation and non-conscious and semi-conscious incubation, to ultimately resolve a solution.

While in reality the four stages of the creative problem solving process are probably not so ordered as the original proponents espoused (Wallas 1925, Hadamard 1945), it is the ability to shift back and forth between the stages, sometimes drawing upon cognitive reasoning, at other times non-cognitive forms of reasoning,(i.e. preparation and incubation) that may prove vital in the creative problem solving process (Russ 1993).

Factor three, which corresponds to the scale for measuring *Spatial/verbal* approaches to reasoning, mirrors the nature of thinking a student may employ be it spatial, verbal, or both spatial and verbal, in seeking a solution to a problem. The negative loading for the statement *thought more with words/symbols than with pictures/diagrams* however, would seem to indicate that the factor is aligned more strongly with the spatial, rather than verbal component of reasoning.

Recent neuro-biological evidence gathered on two systems of reasoning has identified two brain circuits used in mathematical thinking. One circuit is used for approximate arithmetic, the other for exact arithmetic. Interestingly the circuit associated with approximate arithmetic is located in a region strong in visual-spatial processing while that for exact arithmetic is located in a region strong in linguistic processing (Dehaene et al 1999). Thus this factor may be tapping into both visual spatial and linguistic circuits of reasoning.

Factor four, which has been identified as the *Free-flowing* approach to reasoning, mirrors the intent of a student who has employed non-conscious or semi-conscious forms of reasoning, to solve a problem such as that undertaken during incubation. Since the very action of non-conscious reasoning is un-conscious, the only indication that such a form of reasoning has taken place is by a positive response to some indicative behaviours which may reflect the process, such as thinking of the answer while doing something else, or arriving at an answer upon first waking in the morning. In addition, an individual who draws upon this approach to reasoning may well access, stored information, which is more global in its origin.

Of interest is the notion proposed by Torrance and Rockstein (1988) that the inclusion of incubation time during problem solving is critical to the putting together of processes, both

cognitive and non-cognitive, needed for whole brain functioning and the creative solution of problems.

Factor five, which corresponds to the scale for measuring *Feeling* approaches to reasoning, reflects the individual who has followed a feeling or hunch about what to do and tries to verify or check what they have done. In addition a feeling about the quality and correctness of the solution may be present before the answer is checked. Individuals who utilize this approach to reasoning may be synthesizing information drawn from non-conscious forms of reasoning with that of conscious deliberate reasoning.

Interestingly in trying to clarify the difference between intelligence and creativity Sternberg and O'Hara (1999) cite Shouksmith (1973) as making a distinction between the rightness and goodness of a response. In Shouksmith's view judging the rightness of a response is an attempt to measure logical reasoning while judging the goodness of a response is a measure of creativity. The intersection of these two would yield responses that are both right and good.

Accessing Two Underlying Dimensions

On another level of analysis it is also possible to investigate the structure of the five factors using a system of *a priori* classification, undertaken prior to statistical appraisal as a means of identifying a statement as being rule-based or associative in its approach to reasoning. Viewing the labels associated with each of the items permits access to this system of *a priori* classification for additional analysis.

Items that were originally classified as demonstrating rule-based reasoning as indicated by their description of behaviour have the prefix R in front of the item number. Those items that were originally classified as demonstrating associative reasoning as indicated by their description of behaviour have the prefix A in front of the item number. Those items representing both rule-based and associative reasoning have the letters AR or RA as a prefix to the item number. Hence an examination of the factors with this classification in mind yields some interesting findings. A break down of the number and nature of items by factor is given in Table 5.

Table 5. Table showing composition of each factor using a priori *classification of items.*

				Anticipated	Form of	Reasoning
Factor No.	Scale Name	Nos. of Items	Nos. of items classified Rule-based	Nos. of items classified Associative	Nos. of items classified both Rule-based & Associative	Dominant classification
One	Systematic	12	9	1	1	Rule-based
Two	Strategic	8	6	2	0	Rule-based
Three	Spatial/verbal	7	1	3	3	Both /Assoc
Four	Free Flowing	6	0	6	0	Associative
Five	Feeling	10	2	7	1	Associative
Total		43	18	19	5	

The *Systematic* factor and *Strategic* factor appear to be comprised of items that were classified largely as rule-based in their approach to reasoning. The *Free-flowing* factor and the *Feeling* factor appear to be comprised of items that were classified largely as associative in their approach to reasoning. Interestingly the *Spatial/verbal* factor comprises items that were classified as involving both rule-based and associative forms of reasoning while at the same time having a number of statements that were classified as being associative. This would appear to be consistent with the tendency of the factor to be more closely aligned with the spatial rather than verbal component of reasoning.

While this *a priori* method of classifying items may not be excessively generalized to indicate the nature of the underlying dimension within each scale without a good deal of further testing, the evidence accumulated thus far is compelling in its consistency.

CONCLUSION

In conclusion it would appear that cognitive and non-cognitive forms of reasoning may be identified, described and measured in the novel mathematics problem solving context through the use of a self report questionnaire (viz: Systems of Reasoning Questionnaire SRQ). Sloman's (1996) characterization of two forms of reasoning, together with the four stage model of creative problem solving have provided a useful framework by which this can be done. Indeed cognitive and non-cognitive measurements gathered from preliminary data indicate that success in creative problem solving is significantly correlated with the *Feeling* approach to reasoning (Aldous 2001, in press).

While the preliminary work outlined in this paper needs to be replicated with larger samples and tested with additional procedures (e.g. confirmatory factor analysis) the potential for accessing other elements of the teaching and learning system emanates from the administration of an instrument such as this to students undertaking a novel problem solving learning activity. As Biggs (2001) has pointed out, it is the interaction between the student, the task and the context that is vital in shaping future directions in learning and teaching and the students' approaches to reasoning and learning are "the barometers that tell how well the general system is working".

REFERENCES

Aldous, C. (1999). "I used all of me": Listening to the Self in Learning. Paper presented at the Educational research Conference of the Flinders University Institute of International Education, The Flinders University School of Education and The Graduate School of Education, University of Adelaide, Adelaide Australia.

Aldous, C. (2000). Creativity in Problem Solving:Uncovering Cognitive and Non-cognitive Systems in the Solving of Novel Mathematics Problems.Unpublished manuscript, Adelaide Australia.

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: a measure of intuition-analysis for organisational research. Journal of Management Studies, 1, 119-135.

Anderson, J. R. (1997). Cognitive Psychology and its Implications (4th ed.). New York: W.H. Freeman and Company.

Andrich, D. & Luo, G. (1993). A hyperboloc cosine latent trait model for unfolding dichotomous single-stimulus responses. Applied Psycgological Measurement, 17, 253-276.

Andrich, D. (1996). A hyperbolic cosine latent trait model for unfolding polychotmous responses: Reconciling Thurstone and Likert methodologies. British Journal of Mathematical and Statistical Psychology, 49, 347-365.

Baer, J. (1993). Creativity and Divergent Thinking A Task - Specific Approach. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Begg, A. (1999, 4-7 July). Enactivism and Mathematics Education. Paper presented at the Making the Difference Proceedings of the Twenty-second Annual Conference of The Mathematics Education Research Group of Australasia Incorporated, Adelaide, South Australia.

Biggs, J. (1993). From theory to practice:A cognitive systems approach. Higher Education Research and Development, 12, 73-86.

Biggs, J. (2001). Enhancing learnng: A matter of style or approach. In R. J. S. a. L. S. Zhang (Ed.), Perspectives on thinking, learning and cognition (pp. in press): Erlbaum.

Bruner, J. (1986). Actual minds, possible worlds. Cambridge, MA: Harvard University Press.

Bruner, J. (1990). Acts of Meaning. Cambridge, MA: Harvard University Press.

Bryant, F. B. & Yarnold, P. R. (1998). Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis. In L. G. G.rimm & P. R. Yarnold (Eds.), Reading and Understanding Multivariate Statisitics (pp. 99-136). Washington, DC: American Psychological Association.

Carlson, R. & Bailey, J. (1997). Slowing down to the speed of life. Rydalmere, NSW: Hodder and Stoughton.

Carlson, R. (1998). You can be happy no matter what. Moorebank, NSW: A Bantam Book Trans World Publishers (Aust) Pty Limited.

Carroll, J. B. (1983). Studying individual differences in cognitive abilities:Through and beyond factor analysis. In R. F. Dillon. & R. R. Schmeck (Eds.), Individual Differences in Cognition (Vol. 1, pp. 1-33). New York: Academic Press.

Cattell, R. B. (1966). The Scree test for the number of factors. Multivariate Behavioural Research, 1(2), 245-276.

Davis, B. (1996). Teaching Mathematics Toward a Sound Alternative. New York: Garland Publishing.

Davis, B. A., Sumara, D. J. and Kieren, T. E. (1996). Cognition, co-emergence, curriculum. Journal of Curriculum Studies, 18(2), 151-169.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking :Behavioural and Brain-Imaging Evidence. Science Magazine, 284, 970-977.

Getzels, J. W. (1975). Problem finding and the inventiveness of solutions. Journal of Creative Behaviour, 9, 12-18.

Good, R. G., Wandersee, J. & St Julien, J. (1993). Cautionary Notes on the Appeal of Constructivism in Science Education. In K. Tobin (Ed.), The Practice of Constructivism in Science and Mathematics Education (pp. 71-90). Washington, DC: AAAS.

Gunn, S. (1999, 4-7 July). Emerging Themes in Statistics Education. Paper presented at the Making the Difference Proceedings of the Twenty-second Annual Conference of The Mathematics Education Research Group of Australasia , Incorporated, Adelaide, South Australia.

Gustafsson, J. E. & Balke, G. (1993). General and Specific Abilities as Predictors of School Achievement. Multivariate Behavioural Research, 28(4), 407-434.

Hadamard, J. (1945). An Essay on the Psychology of Invention in the Mathematical Field. Princeton, N.J.: Princeton University Press.

James, W. (1950). The principles of psychology. New York: Dover.

Johnson-Laird, P. N. (1983). Mental Models. Cambridge, MA: Harvard University Press.

Kaiser, H. F. & Caffrey, J. (1965). Alpha factor analysis. Psychometrika, 30(1), 1-14.

Keeves, J. P. (1966). Students' Attitudes Concerning Mathematics A Study of The Opinions and Views Concerning Mathematics and Mathematics Teaching Held by Students in the Upper Primary and Lower Secondary Schools of New South Wales and Victoria. Unpublished Degree of Master of Education, University of Melbourne, Melbourne.

Keeves, J. P. (1997). Multivariate Analysis. In J. P. Keeves (Ed.), Educational Research, Methodology, and Measurement (second ed., pp. 403 - 412). Cambridge, UK.: Cambridge University Press.

Kline, P. (1993). The Handbook of Psychological Testing. London, UK: Routledge.

Koberg, D. a. B., J. (1976). The All New Universal Traveller: a soft-systems guide to creativity, problem solving and the process of reaching goals.

Lovitt, C. (1999). Investigations as a central focus for a mathematics curriculum. Paper presented at the Proceedings of the second Australian Association of Mathematics Teachers Virtual Conference.

Miller, G. A., & Johnson-Laird, P. N. (1976). Language and Perception. Cambridge, Mass: The Belknap Press of Harvard University Press.

Neisser, U. (1963). The multiplicity of thought. British Journal of Psychology, 54, 1-14.

Parr, G. P. (1974). The Phenomenon of Mathematical Invention. Mathematics Teaching, 66, 56-58.

Phillips, D. C. (2000). Constructivism in Education: Opinions and Second Opinions. Chicago: University of Chicago Press.

Piaget, J. (1926). The language and thought of the child. London: Routledge and Kegan.

Runco, M. A. & S.O., S. (1999). Experimental Studies of Creativity. In R. J. Sternberg (Ed.), Handbook of Creativity (pp. 62-92). New York , N. Y.: Cambridge University Press.

Russ, S. (1993). Affect and Creativity: The Role of Affect and Play in the Creative Process. Hilldale, New Jersey: Lawrence Erlbaum Associates.

Shouksmith, G. (1973). Intelligence, creativity and cognitive style. London: Angus and Robertson.

Shuell, T. J. (1986). Cognitive conceptions of learning. Review of Educational Research, 26(3), 4-15.

Sloman, S. A. (1996). The Empirical Case for Two Systems of Reasoning. Psychological Bulletin, 119(1), 3-22.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P., & Grigorenko, E. L. (1996). Identification, instruction and assessment of gifted children: A construct validation of a triarchic model. Gifted Child Quarterly, 40, 129-137.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998). Teaching Triarchically Improves School Achievement. Journal of Educational Psychology, 99(3), 374-384.

Sternberg, R. J. & O'Hara, L. A. (1999). Creativity and Intelligence. In S. R. J. (Ed.), Handbook of Creativity (pp. 251-272). New York, N.Y.: Cambridge University Press.

Taylor, C. W. (1989). Various Approaches to and Definitions of Creativity. In R. J. Sternberg (Ed.), The Nature of Creativity: Contemporary Psychological Perspectives (pp. 99-121). New York, N.Y.: Cambridge University Press.

Taylor, P. C. (1996). Mythmaking and myth breaking in the mathematics classroom. Educational Studies in Mathematics, 3, 1521-1573.

Torrance, E. P. a., & Rockenstein, Z. L. (1988). Styles of Thinking and Creativity. In R. R. Schmeck (Ed.), Learning Strategies and Learning Styles (pp. 275-290). New York, N.Y.: Plenum Press.

Vygotsky, L. S. (1987). Thinking and speech. In R. W. R. a. A.S.Carton (Ed.), The collected works of L.S. Vygotsky. Problems of general psychology (Vol. 1). New York: Plenum Press.

Wallas, C. (1926). The art of thought. New York: Harcourt brace.

Zajonic, R. B. (1980). Feeling and Thinking Preferences Needs No Inferences. American Psychologist, 35(2), 151-175.

APPENDIX 1

Table 5. Principal component solution rotated to orthogonal structure derived from Space data N=100 (Loadings >= 0.3). Items in common with Number are indicated with #.

Item	Common	Statement	1	2	3
		Free Flowing Factor			
R32		I used a strategy/procedure I learned in class to solve this problem.	0.57		
A17		I got the answer but then had some trouble explaining how I arrived at it.	0.55		
A40	#	I thought of how to get the answer to this problem when I woke up in the morning.	0.54		
A37	#	I tried to be inventive in solving this problem.	0.54		
A36	#	I drew upon all mental resources/parts of me to solve this problem.	0.47		
A39	#	I thought of how to get the answer to this problem while I was doing something else (e.g. sitting in school, working another problem, riding a bike).	0.46		
A23	#	I associated this problem with things I experienced/learned outside of school..	0.43		
R21		I used things I learned in school while solving this problem	0.421		0.33
A35		I let my mind go free, thinking of any possibility when solving this problem.	0.42	-0.36	
A41	#	I tried a lot of different ways until I found the right one.	0.407		
		Systematic Factor			
A28		I had a sense of number size in this problem.		0.74	
R26		I recognised number patterns in solving this problem.		0.61	
R33	#	I used specific formulae in solving this problem.		0.56	
R16	#	I got the answer by using mathematical steps right from the beginning.		0.53	
A24		I associated this problem with things/images/pictures/diagrams etc. I have seen.		0.44	
R30	#	I recalled specific maths facts in solving this problem	0.39	0.44	
R29	#	I organised my reasoning/thinking in a strategic way.		0.40	0.31
		Feeling Factor			
R34	#	I tried to verify/check that my answer was correct.			0.56
A38	#	I developed a feeling about the correctness of my solution before I checked it.	-0.35	0.41	0.51
A19		I used my imagination in solving this problem		-0.49	0.51
A25		I saw spatial/visual patterns in my mind			0.50
A18	#	I just knew how to get the answer without using mathematical reasoning first.			0.46
A20	#	I used my intuition in solving this problem			0.44
R31		I used a sequence of logical steps in this problem			0.43
A27	#	I felt that I related to/connected with the patterns in this problem.			0.40
A22	#	The “Answer” or “How to get the answer” suddenly came into my head while I was working on this problem.			0.39

*Table 6. Table showing summary details of the three factored scales based on **Space** data*

Factor No.	Scale Name	Nos. of Items	Nos. of Common Items	Reliability Kaiser-Caffrey	% Variance explained before rotation	% Variance explained after rotation	Dominant classif-ication
One	Free-flowing	10	6	0.67	13.27	10.73	Associative
Two	Systematic	7	4	0.69	9.00	10.58	Rule-based
Three	Feeling	9	5	0.60	8.48	9.44	Associative

APPENDIX 2

*Table 7. Principal component solution rotated to orthogonal structure derived from **Number** data N=102. (Loadings >= 0.3). Items in common with Space are indicated with #.*

Item	Common	Statement	1	2	3
		Systematic Factor			
R16	#	I got the answer by using mathematical steps right from the beginning.	0.64		
R31		I used a sequence of logical steps in this problem.	0.64		
R29	#	I organized my reasoning/thinking in a strategic way.	0.61		.0.40
R30	#	I recalled specific maths facts in solving this problem.	0.60	0.34	
R32		I used a strategy/procedure I learned in class to solve this problem.	0.52	0.45	
A35		I let my mind go free, thinking of any possibility when solving this problem.	-0.52	0.34	
A17		I got the answer but then had trouble explaining how I arrived at it.	-0.49		
R21		I used things I learned in school while solving this problem.	0.48	0.32	
R33	#	I used specific formulae in solving this problem	0.47		
		Free-flowing Factor			
A39	#	I thought of how to get the answer to this problem while I was doing something else (e.g. sitting in school, working another problem, riding a bike)		0.58	
A23	#	I associated this problem with things I experienced/learned outside of school.		0.57	
A37	#	I tried to be inventive in solving this problem.		0.54	0.36
A24		I associated this problem with things/images/pictures/diagrams etc. I have seen.		0.53	
A40	#	I thought of how to get the answer to this problem when I woke up in the morning.		0.53	
A36	#	I drew upon all mental resources/parts of me to solve this problem.		0.50	
A25		I saw spatial/visual patterns in my mind.		0.48	
A19		I used my imagination in solving this problem	-0.39	0.48	
A41	#	I tried a lot of different ways until I found the right one.		0.35	
		Feeling Factor			
R34	#	I tried to verify/check that my answer was correct.			0.66
A28		I had a sense of number size in this problem			0.66
AR15		I got the answer first and then thought out the mathematical reasons for it being correct.			0.56
A38	#	I developed a feeling about the correctness of my solution before I checked it.			0.55
R26		I recognized number patterns in solving this problem.	0.30		0.54
A18	#	I just knew how to get the answer without using mathematical reasoning first.			0.44
A22	#	The “Answer” or “How to get the answer” suddenly came into my head while I was working on this problem.			0.44
A27	#	I felt that I related to/connected with the patterns in this problem.			0.41
A20	#	I used my intuition (followed a gut feeling) solving this problem		0.299	0.29

*Table 8. Table showing summary details of the three factored scales based on **number** data*

Factor No.	Scale Name	Nos. of Items	Nos. of Common Items	Reliability Kaiser-Caffrey	% Variance explained before rotation	% Variance explained after rotation	Dominant classi-fication
One	Systematic	9	4	0.76	16.23	12.12	Rule-based
Two	Free-flowing	9	6	0.69	11.13	11.75	Associative
Three	Feeling	8	5	0.68	7.56	11.11	Associative

Researching computer self-efficacy

Olivia Khorrami-Arani
Flinders University, School of Education *khor0012@flinders.edu.au*

The purpose of this paper is to develop a measure of computer self-efficacy, administer this measure to obtain baseline data at the start of a course on Information Technology (IT), and to describe the IT course for Year 8 in one secondary school. An online questionnaire administered at the beginning of the IT semester-length course, gathered information pertaining to the students' background, prior computer experience, comfortableness with using computers, and their attitudes towards computers. During the course, observations were made in the six Year 8 IT classrooms under study. This paper reports a work in progress. The principal focus is on the selection and development of research instruments, the administration of the online questionnaire, the analysis of the student data at the start of the course, and the description of the classroom observations. Preliminary analysis suggests the measure of computer self-efficacy adopted in the study is an effective tool.

Computer self-efficacy, computer attitudes, online questionnaire, information technology

Computers have made a dramatic impact on our society, particularly in the field of education. Computers are common tools in most schools, and are being used increasingly in all subject areas. Although some students are enthusiastic about using computers, others may be more apprehensive. In so far as computers aid learning and are common tools in the workforce, it is crucial for all students to become familiar and comfortable with their use. According to Ertmer et al. (1994), successful computer experiences prepare students to participate effectively in a computer-dominated society. Computer experience alone is not the only precursor to student success with computers. Computer knowledge and attitudes also play an important role. Attitudes towards computer technologies are associated with a concept known as computer self-efficacy (Delcourt and Kenzie, 1993), which, in turn, has proven to be a factor in understanding the frequency and success with which individuals use computers (Bandura, 1986, Compeau and Higgins 1995).

The purpose of this paper is to develop a measure of computer self-efficacy to obtain baseline data for a wider study of computer self-efficacy in Year 8 IT classes. Although this paper only provides some initial findings, the aim of the broader study is to gauge growth in computer self-efficacy over an IT course for Year 8 students.

THE CONCEPT AND MEASUREMENT OF SELF-EFFICACY

Bandura (1977) was the first writer to use the term self-efficacy. The concept of self-efficacy and computer self-efficacy are discussed. In addition, the most widely used measures of computer self-efficacy are presented.

THE CONCEPT OF SELF-EFFICACY

Bandura (1986) defined self-efficacy as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances. It is concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses" (p. 391). Therefore, the concept of self-efficacy was context specific, or the valuing of self through specifically defined situations. The definition of self-efficacy

provided by Bandura (1986) highlighted the importance of distinguishing between component skills and the ability to perform actions. Further studies by Bandura discussed the psychological construct of self-efficacy as a concept that referred "to beliefs in one's capabilities to mobilize the motivation, cognitive resources and courses of action needed to meet situational demands" (Bandura and Wood, 1989, p. 506). In addition, self-efficacy influenced people's decisions, goals, their amount of effort in conducting a task, and the length of time they would persevere through obstacles and difficulties.

Oliver and Shapiro (1993) have described the importance of not confusing self-efficacy with self-esteem. Although self-efficacy and self-esteem are both concerned with the judgments of individuals, they do not share any direct relationship with each other. They noted that self-efficacy was concerned with self-capability, whereas self-esteem was concerned with self-worth.

Computer Self-efficacy

The term *self-efficacy* was soon extended to particular domains, including the use of computers. Compeau and Higgins (1995) defined computer self-efficacy as "a judgment of one's capability to use a computer" (p. 192). It was noted that self-efficacy judgments could influence an individual's expectations because "the outcomes one expects derive largely from judgments as to how well one can execute the requisite behaviour" (Bandura, 1978, p. 241). Computer self-efficacy has a major impact on an individual's expectations towards using computers according to Compeau and Higgins (1995). In addition, individuals who did not see themselves as competent computer users were less likely to use computers (Kinzie and Delcourt, 1991; Oliver and Shapiro, 1993).

Studies of computer self-efficacy have been conducted on individuals in the work force (Gist. et al, 1989; Burkhardt and Brass, 1990; Compeau and Higgins, 1995; Harrison and Rainer, 1997; Decker, n.d). These studies demonstrated the impact that computer self-efficacy has on increasing performance and the technological innovation of employees, reducing computer-induced anxiety, and promoting higher occupational positions. Other computer self-efficacy studies have used student subjects at a university level (Karsten and Roth, 1998a; 1998b; Langford and Reeves, 1998). Overall, these studies showed that higher levels of computer self-efficacy corresponded to increased performance in computer courses and a greater achievement of computer competency.

A recent study conducted by Wallace (1999) investigated and described four main factors that influenced the development of computer self-efficacy. These factors were computer anxiety, computer confidence, computer liking and computer knowledge. He demonstrated a significant correlation between the computer self-efficacy model (composed of the four mentioned factors) and a 3-item measure of computer self-efficacy, further confirming that the model was a valid means to explore components of the computer self-efficacy construct. Comparisons were made between the computer self-efficacy of education and computing students. Wallace (1999) reported that computing students expressed low levels of computer anxiety, and higher levels of computer knowledge, computer liking, and computer confidence in comparison with education students.

Measurements of Computer Self-efficacy

Several research studies have investigated computer self-efficacy, using a variety of measures to study different subject groups. The computer self-efficacy scale created by Murphy et al. (1989) is a popular measure used by many researchers in the field of technology. The Murphy (1989) scale (see Table 1) was developed for measuring individuals' self-perceptions of accomplishments surrounding particular computer-related knowledge and skills. The 32-item scale is composed of three different levels of computing skills: beginner's level, advanced level, and a level associated with mainframe computers.

Table 1. The Murphy (1989) Computer Self-efficacy Scale

Level of Computer Skills	Item Number	Statement
Beginning	1	I feel confident working on a personal computer (microcomputer)
	2	I feel confident getting the software up and running
	6	I feel confident entering and saving data (numbers or words) into a file
	7	I feel confident escaping/exiting from a program or software
	9	I feel confident choosing a data file to view on a monitor screen
	12	I feel confident handling a floppy disk correctly
	15	I feel confident making selections from an onscreen menu
	17	I feel confident using a printer to make a "hard copy" of my work
	18	I feel confident copying a disk
	19	I feel confident coping an individual file
	20	I feel confident adding and deleting information to and from a data file
	21	I feel confident moving the cursor around the monitor screen
	23	I feel confident using the computer to write a letter or essay
	27	I feel confident storing software correctly
	30	I feel confident getting rid of files when they are no longer needed
	31	I feel confident organizing and managing files
Advanced	5	I feel confident using the user's guide when help is needed
	10	I feel confident understanding terms/words relating to computer hardware
	11	I feel confident understanding terms/words relating to computer software
	13	I feel confident learning to use a variety of programs (software)
	14	I feel confident learning advanced skills within a specific program (software)
	16	I feel confident using the computer to analyze number data
	22	I feel confident writing simple programs for the computer
	24	I feel confident describing the function of computer hardware (keyboard, monitor, disk drives, processing unit)
	25	I feel confident understanding the three stages of data processing: input, processing, output
	26	I feel confident getting help for problems in the computer system
	28	I feel confident explaining why a program (software) will or will not run on
	29	I feel confident using the computer to organize information
	32	I feel confident troubleshooting computer problems
Mainframe	3	I feel confident logging onto a mainframe computer system
	4	I feel confident working on a mainframe computer
	8	I feel confident logging off the mainframe computer system

Many researchers have adapted the original Murphy computer self-efficacy scale (Langford and Reeves, 1998; Davis and Davis, 1990; Harrison and Rainer, 1992; Harrison and Rainer, 1997; Decker, n.d.), while others have adapted a slightly modified version of the Murphy scale for their study (Karsten and Roth, 1998a; Karsten and Roth, 1998b; Torkzadeh and Koufteros, 1994; Delcourt and Kinzie, 1993; Ertmer et al., 1994; Qutami and Abu-Jaber, 1997; Zhang and Espinoza, 1998). Some researchers have developed their own measure of computer self-efficacy because they have found previous scales inadequate (Gist et al., 1989; Burkhardt and Brass, 1990; Compeau and Higgins, 1995; Havard and Atkinson, n.d; Hill et al., 1987; Miura, 1987; Vasil et al., 1987; Wallace, 1999 Webster and Martocchio, 1992).

RESEARCH METHODS

The setting of the study and the participants involved are described in this section. The research design is detailed along with several test measures used to investigate computer self-efficacy in IT classes.

Setting of the Study

Wallace (1999) suggested that the study of computer self-efficacy was lacking in school classrooms and that such studies in specific subject classrooms would be beneficial. Accordingly, the present study was conducted in one secondary school in Adelaide, South Australia, a public co-education school located in a medium to high socio-economic area. This secondary school has a long-term commitment, and involvement with the use of computers as teaching and learning tools. The school has an established Information Technology program conducted by dedicated teachers who were willing to aid in this investigation.

Selection of Sample

The participants of this study are enrolled in six different Year 8 IT classes in the one school. There are 105 students, 61 females and 44 males.

Research Design

The design of this study is to obtain baseline data at the beginning of a semester course in IT. The IT course comprises four units: word-processing, drawing, spreadsheets and databases. Microsoft Word and Excel were the programs used for teaching purposes and instructional manuals were created for each of the four units. Slight modifications were made to each of the instructional booklets by some of the teachers. The method of assessment is identical in all the classes observed and includes various assignments and tests. To obtain an initial level of performance, a measure was administered at the start of the course. Observations were also recorded on the basis of identifying any differences between the six IT classes.

Test Measures

The questionnaire is divided into three parts. Part 1 asks students to provide background information about themselves and any prior computer experience they may have had. Part 2 is a measure of computer self-efficacy. Because the term *computer self-efficacy* would be foreign to Year 8 students, the phrase, *Feeling Comfortable with Computers*, is used instead in the questionnaire. Part 3 aims to assess students' attitudes towards computers. Another component of the test measures were any classroom observations made during the study.

Student Background

Each subject was given a number corresponding to their classroom-seating plan, so their names remain anonymous. Since all results are to be completely confidential, these student numbers are not used for any other purpose than linking students to their questionnaires. Participants were also asked to indicate their gender, prior use of, and access to computers.

Computer Self-efficacy

To measure computer self-efficacy, the Murphy et al. (1989) scale was used. This scale is adjusted to suit the comprehension of Year 8 Information Technology students. Two items from the original scale were omitted because they were not relevant to computer activities for year 8 IT classes. The items in this study were reordered according to logically ordered scale used by Karsten and Roth, 1998a and 1998b. These statements are laid out in an easily understood and systematic manner, ideal for Year 8 students. The modification and reordering of items is detailed in Table 2.

Computer Attitudes

The computer attitude scale comprises three components: computer anxiety, computer confidence, and computer liking. This computer attitude scale is based on the Loyd and Gressard's original format (1984). These scales are quite widely used (for example, by Massoud, 1991 and Wallace, 1999). A revised format of the computer attitude scale was

necessary to suit the subjects involved in this study who are at the beginning of secondary school. The modification of items is detailed in Table 3.

Table 2. Original and Modified Items in the Murphy (1989) Scale

Statement Number	Original Statement	Modified Statement
1	Entering and saving data (words and numbers) into a file	Entering and saving words or numbers into a file
2	Calling up a file to view on a monitor screen	Choosing a file to view on a monitor screen
4	Escaping/exiting from a program or software	Escaping/exiting from a program
5	Making selections from an on-screen menu	Choosing items from an on-screen menu
6	Copying an individual file	Copying a file
10	Using a printer to make a "hard copy" of my work	Using a printer
11	Getting rid of files when they are no longer needed	Getting rid of files on a computer
13	Adding and deleting information to and from a data file	Adding and deleting information to and from a file
18	Describing the function of computer hardware (keyboard, monitor, disk drives, processing unit)	Describing what computer hardware does (keyboard, monitor, disk drives, processing unit)
19	Troubleshooting computer programs	Solving computer problems
20	Explaining why a program (software) will or will run on a given computer	Explaining why a computer program does not work on a computer
22	Learning to use a variety of programs (software)	Learning to use different computer programs
24	Learning advanced skills within a specific program (software)	Learning skills to use in a computer program
27	Getting help for problems in the computer system	Getting help for problems when using a computer
	Storing software correctly	(This item was omitted)
	Writing simple programs for the computer	(This item was omitted)

Table 3. Original and Modified Items in the Three Sub-scales of Computer Attitudes

Attitude Sub-scale	Statement Number	Original Statement	Modified Statement
Computer Anxiety	4	I feel aggressive and hostile towards computers	I feel angry towards computers
Computer Confidence	10	I have a lot of self-confidence when it comes to working with computers	I feel confident when it comes to working with computers
Computer Liking	2	The challenge of solving problems with computers does not appeal to me	I do not like problems with computers
	3	I think working with computers would be enjoyable and stimulating	I think working with computers would be fun
	5	When there is a problem with the computer program that I can't immediately solve, I would stick with it until I have the answer	I would quit trying to solve a computer problem
	9	If a problem is left unsolved in a computer case, I would continue to think about it afterwards	I think about computer problems when they are left unsolved

Online Measures

All test measures consisting of students' background information, computer self-efficacy and computer attitudes were developed in online format, and were administered during classroom time. The online questionnaire was administered during a double lesson in each IT class.

Classroom Observations

During the semester long IT course observations were made in the six Year 8 IT classes. Six different teachers taught these classes. Observations were used as a means of elaborating the learning activities in each class according to teaching styles and classroom dynamics.

ANALYSIS

In the study reported here, only Term 3 of the IT classes was observed. The term consisted of 6 weeks of word-processing and 3 weeks of computer drawing. The number of students in each of the six classes and the teachers' level of experience is detailed in Table 4.

Table 4. A comparison between students and teachers in six IT classrooms

Class No.	No. students	No. questionnaire participants	Teacher's level of experience	E-mail taught by the teacher
I	16 (6 M, 10 F)	15 (6 M, 9 F)	First year	Yes
II	20 (10 M, 10 F)	18 (9 M, 9 F)	First year	Yes
III	17 (7 M, 10 F)	15 (5 M, 10 F)	Over 20 years	No
IV	20 (11 M, 9 F)	20 (11 M, 9 F)	16 years	No
V	23 (11 M, 12 F)	18 (7 M, 11 F)	15 years	No
VI	19 (7 M, 12 F)	19 (7 M, 12 F)	7 years	No

The teachers in both classes I and II had never taught Year 8 IT prior to this class. The instruction of email was semester long in classes I and II but was not taught as a component of the course in Term 3 of classes III, IV, V and VI.

The number of students who completed the questionnaire at the beginning of the IT course is indicated in Table 4. This was a response rate of 90 percent.

The reliability of the three attitude sub-scales and of the measure of computer self-efficacy was computed using coefficient alpha. The results together with the number of items in each computer scale are presented in Table 5. The reliability of all measures, and particularly of the computer self-efficacy scale, shows a high level of internal consistency.

Table 5. Reliability estimates (coefficient alpha) and the number of items in four computer scales (N=105)

Computer Scale	Coefficient Alpha	Number of items
Computer Anxiety	0.77	10
Computer Confidence	0.88	10
Computer Liking	0.85	10
Computer Self-efficacy	0.95*	30

* based on 104 subjects because of an omitted item by one student

A measure of the concurrent validity of the four computer scales adopted in this study is obtained by Pearson's product-moment correlation, the results of which are displayed in Table 6. All intercorrelation coefficients are significant at the 0.05 level (103 df).

Table 6. Intercorrelation coefficients for four computer scales

	Computer Anxiety	Computer Confidence	Computer Liking	Computer Self-efficacy
Computer Anxiety	1.00			
Computer Confidence	0.74*	1.00		
Computer Liking	0.63*	0.77*	1.00	
Computer Self-efficacy	0.63*	0.78*	0.56*	1.00

* significant at 0.05 level (103 df)

Students' attitudes towards computers (anxiety, confidence, and liking), their feeling of comfortableness about computers (computer self-efficacy), and their perceived levels of

performance in different aspects of computing are shown in the means in Table 7. At the start of the IT course, students have little anxiety towards computers ($\bar{x}$ = 4.2) and are quite confident ($\bar{x}$ = 3.7), but are equivocal as far as liking computers is concerned ($\bar{x}$ = 3.1). Their comfortableness about their ability to use computers (computer self-efficacy) is also quite high at the start of the course ($\bar{x}$ = 4.1). These four measures have a maximum of 5. Their previous computer use ($\bar{x}$ = 4.5, maximum of 5) and computer experience ($\bar{x}$ = 2.6, maximum of 4) is also quite high. Students' stated use with computer games, word processing, the Internet, and email is quite high with a mean of approximately 1.1 (out of 2), but they say they have little use of graphics, spreadsheets or databases (see Table 7).

Table 7. Means and standard deviations for students' attitudes towards computers, computer self-efficacy, experience, use, and perceived levels of performance in computing

	Mean	Maximum	Standard Deviation
Computer Attitudes			
Anxiety*	4.2	5	0.68
Confidence	3.7	5	0.88
Liking	3.1	5	0.89
Computer Self-efficacy	4.1	5	0.64
Computer Experience	2.6	4	0.95
Computer Use	4.3	5	0.81
Levels of Use			
Computer Games	1.3	2	0.76
Word-processing	1.1	2	0.55
Internet	1.2	2	0.66
Email	0.9	2	0.74
Graphics	0.4	2	0.64
Spreadsheets	0.3	2	0.44
Databases	0.2	2	0.39

* This scale was negatively scored such that a high mean indicated little anxiety towards computers.

Space here does not allow for elaboration of the classroom observations or the detailed content of the IT course in use at the school. Perhaps, however, the classroom situations depicted in Figures 1 and 2 convey something of the enjoyment of students in different computing activities during the course.

Figure 1. Year 8 IT students working on a class assignment

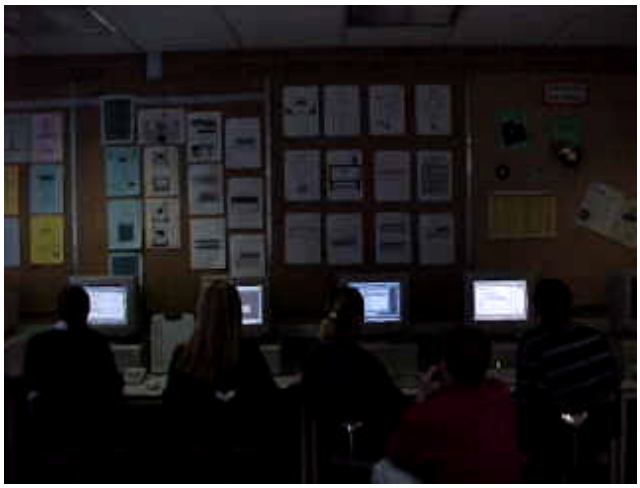


Figure 2. Teacher aiding a student

CONCLUSION

In Wallace's (1999) major study of computer self-efficacy, he observed that there were very few studies at the school level, and he went on to suggest that it would be useful to conduct studies of computer self-efficacy in specific subject classrooms. This paper outlines the beginning of a longitudinal study over the duration of course in IT at the lower secondary school level.

The purposes of the study reported here were achieved in the modification and adoption of a measure of computer self-efficacy and three computer attitude scales. All four measures were appropriate for students at the Year 8 level in terms of comprehensibility. Importantly for the planned longitudinal study, the measure of computer self-efficacy is highly reliable, while the three sub-scales of computer attitudes have quite satisfactory reliabilities. As expected, the measures correlate with each other quite highly. Students' attitudes towards computers and their degree of confidence in their use of computers are quite high. It is for the longitudinal study, of which this study is a part, to gauge the growth in students' attitudes and computer self-efficacy and to attempt to explain change in terms of teaching styles and classroom dynamics.

REFERENCES

Bandura, A. (1977) Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review,* 84 (2), 191-215.

Bandura, A. (1986) *Social Foundations of Thought and Action*. NJ: Prentice Hall.

Bandura, A. and Wood R. (1989) Effect of perceived controllability and performance standards on self-regulation of complex decision-making. *Journal of Personality and Social Psychology*, 56 (5), 805-814.

Burkhardt, M.E., and Brass, D.J. (1990) Changing patterns or patterns of change: The effects of a change in technology on social network structure and power. *Administrative Science Quarterly*, 35, 104-127.

Compeau, D.R. and Higgins, C.A. (1995) Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19 (2), 189-212.

Davis, L.D., and Davis, D.F. (1990) The effect of training techniques and personal characteristics on training end users of information systems. *Journal of Management Information Systems,* 7 (2), 93-110.

Decker, C.A. (n.d) Training transfer: Perceptions of computer use self-efficacy among university employees. *Journal of Vocational and Technical Education*. Vol.14 (2), 1-14. [Online] http://scholar.lib.vt.edu/ejournal/JVTE/v14n2/JVTE-3.html [2001, June 25].

Delcourt, M.A.B. and Kinzie, M.B. (1993) Computer technologies in teacher education: The measurement of attitudes and self-efficacy. *Journal of Research and Development in Education,* 27 (1), 35-41.

Ertmer, P.A. Evenbeck, E., Cenramo, K.S., and Lehman, J.D. (1994) Enhancing self-efficacy for computer technologies through the use of positive classroom experiences. *Educational Technology Research and Development,* 42 (3), 45.

Gist, M., Shwoerer, C.E., and Rosen, L.D. (1989) Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology,* 74, 884-891.

Harrison, A.W., Rainer, R.K. (1992) The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems,* 9 (1), 93-111.

Harrison, A.W. and Rainer, R.R. (1997) Testing the self-efficacy-performance linkage of social-cognitive theory. *Journal of Social Psychology*, 137 (1), 79-87.

Havard, B.C, and Atkinson, F.D. (n.d) Factors influencing computer self-efficacy in an introductory level computer course. *Design: Connect, Create, Collaborate*. University of Georgia, Department of Instructional Technology, 134-137 [Online] http://itechl.coe.uga.edu/designconference/PDFs/20.pdf [2001, November 6].

Hill, T. Smith, N.D. and Mann, M.F. (1987) Role of efficacy expectations in predicting the decision to use advanced technologies: The case of computers. *Journal of Applied Psychology,* 72 (2), 307-313.

Karsten, R. and Roth, R. (1998a) The relationship of computer experience and computer self-efficacy to performance in introductory computer literacy courses. *Journal of Research on Computing in Education,* 31 (1), 14-22.

Karsten, R. and Roth, R. (1998b) Computer self-efficacy: A practical indicator of student computer competency in introductory IS courses. *Informing Science*, 1 (3),61-68.

Kinzie, M.B. and Delcourt, M.A.B. (1991) *Computer technologies in teacher education: The measurement of attitudes and self-efficacy*. Paper presented at annual meeting of the American Education Research Association, Chicago. (ERIC Document No. ED 331 891).

Langford, M. and Reeves, T.E. (1998) The relationship between computer self-efficacy and personal characteristics of the beginning information systems student *Journal of Computer Information Systems*, 38 (4), 41-45.

Loyd, B.H. and Gressard, C. (1984) *The effects of sex, age, and computer experience on computer attitudes.* Paper presented at the annual meeting of the Eastern Educational Research Association, Florida.

Massoud, S.L. (1991) Computer attitudes and computer knowledge of adult students. *Journal of Educational Computing Research,* 7 (3), 269-291.

Miura, L. (1987) Gender and socioeconomic stated differences in middle school computer test and use. *Journal of Early Adolescence*, 7, 243-254.

Murphy, C.A., Coover, D., and Owen, S.V. (1989) Development and validation of the computer self-efficacy scale. *Educational and Psychological Measurement,* 49, 893-899.

Oliver, T.A. and Shapiro, F. (1993) Self-efficacy and computers. *Journal of Computer Based Instruction,* 20 (3), 81-85.

Qutami, Y. and Abu-Jager, M. (1997) Student's self-efficacy in computer skills as a function of gender and cognitive learning style at Sultan Qaboos University. *International Journal of Instructional Media,* 24 (1), 63-75.

Torkzadeh, G. and Koufteros, X. (1994) Factorial validity of a computer self-efficacy scale and the impact of computer training. *Educational and Psychological Measurement,* 54 (3), 813-821.

Vasil, L., Hesketh, B. and Podd, J. (1987) Sex differences in computing behaviour among secondary school pupils. *New Zealand Journal of Educational Studies*, 22 (2), 201-214.

Wallace, A.R. (1999) *An exploratory study of the factors influencing the construction of computer self-efficacy*. Unpublished doctoral thesis, Charles Sturt University.

Webster, J. and Martocchio, J.J. (1992) Microcomputer playfulness: Development of a Measure with workplace implications. *MIS Quarterly*, 16 (2), 201-226.

Zhang, Y. and Espinoza, S. (1998) Relationships among computer self-efficacy, attitudes toward computers, and desirability of learning computing skills. *Journal of research on Computing in Education,* 30 (4), 420-438.

Back from the edge: Attitudes towards retrieving a minority language – The case study of Bable (Asturias, Spain) using a survey of teacher attitudes

Lynn M F Arnold
University of Adelaide *arnoldl@wva.org.au*

A survey undertaken of teachers and others attending a conference for the resuscitation of the Asturian language (Bable) held in 1995 in the Principality of Asturias provided an opportunity to explore motivations for promoting this marginal L1 existing in the context of a dominant L2 (Castilian). At the time the author was a visiting scholar at the University of Oviedo and had commenced doctoral studies at the University of Adelaide.

The results of the survey were analysed to determine intentions and opinions of attendees with respect to Bable; and also to analyse less intentional language practices.

The analysis found that there were significant variations in the motivations, knowledge levels and utilisation patterns of those attending. The conclusion is that the formulation of language policy for marginalised languages requires a complex understanding of micro level issues for such policy to be effective.

With respect to less intentional language practices, the analysis found significant syntactical and morphological variation within the attendees' use of Bable. It also found a significant degree of interference between Castilian and Bable.

Key words: Bable, diglossia, interference, normalisation, sociolinguistics

SOCIOLINGUISTIC CONTEXT

Bable is a romance language which surveys indicate is at least understood by nearly 40% of the one million people of Asturias. It has many similarities to Castilian but there are notable lexical and grammatical differences. Though some have argued that it is a dialect of Castilian it should more properly be considered a dialect of Latin whose close relationship with the more dominant Castilian can be classified as a *duolect* or twin relationship.

Bable, which itself has a number of distinct dialects (giving rise to a school of opinion that it is a group of speeches rather than one language), has traditionally been strongest in rural and mining communities. With the exception of the largest city in the region (Xixón/Gijón – an industrial centre and seaport), the urban context has proven very difficult for the language due to status problems.

The utilisation pattern according to domain and genre of Bable and Castilian is diglossic with Bable generally limited to use within domains of close personal relationship (family, friends and certain types of work environment) and less sophisticated genres (tales, festivals, humour &c).

CONFERENCE

Each year since the late 1980s the *Academia de la Llingua Asturiana* has organised conferences to develop capacity in Bable. The conferences have been directed towards teachers but others have also been permitted to attend.

In August 1995, the *Academia* convened a conference entitled *Programa de formación y reciclax* [Training and resuscitation program] which dealt with cultural studies, language improvement and translation issues. One hundred and twenty two people attended this conference. A number had attended similar conferences in previous years.

SURVEY

I sought permission from the president of the *Academia* to conduct a survey of conference attendees. The survey form, written in Castilian, contained the following open-ended questions/requests of conference attendees:

- Write some words by way of self-portrait.
- Why did you attend this conference?
- Do you speak Asturian in your daily life? If so, in what domains?
- Do you believe you believe you speak it well?
- How did you learn the language?
- What is your opinion on (i) normalisation; and (ii) officiality for the language?
- Write ten significant words from the Asturian language which occur to you at this time.

With the exception of the question on normalisation and officiality, none of the questions/requests proposed any possible words which might have prompted certain responses.

The attendees had the following characteristics:

- 54.1% were teachers, 13.9% student teachers with the remainder from a variety of occupations (including miner, lawyer, biologist, journalist and engineer)
- 115 indicated their gender in the survey (84 women and 31 men).
- The average age of those so indicating was 32.2 years (32.6 for women and 31.9 for men). 35.2% of those reporting age were in the 20-29 age group, 27.9% in their thirties and 15.6% forty or more.
- 56.6% indicated they came from the Central zone of the province, 8.2% from the Occidental, 4.9% from the Oriental and 0.8% (one person) from a distinct language area in the far-west of the province. 29.5% did not indicate a zone or came from outside the region.
- 23.8% indicated a connection (either by birthplace, childhood or adult location) that they were connected with an urban location, 16.4% with a mining community, 36.9% a rural zone with the remaining 23% either not reporting any such location or coming from outside the region.

Two of the responses were written in Castilian, the remainder in either Bable or *Asturianu amestao* [*Mixed Asturian* – ie interferred with Castilian]. The responses were analysed both quantitatively and qualitatively.

QUANTITATIVE ANALYSIS

The responses to six of the questions were tabulated firstly as raw data, then as percentages and finally as factors according to the formula below. In each instance the results were tabulated by cross referencing Personal characteristics (gender, occupation, age, language zone and birthplace or current setting) with responses regarding Bable (motivation for attending the conference, language capacity, where knowledge of Bable acquired and domains of usage). In many instances the responses fell into more than one subcategory, hence the combined totals sometimes exceeded 122 (the number of conference attendees).

$$\frac{\mathbf{s}}{\Sigma_{\mathbf{s}}} \div \frac{\mathbf{n}}{\Sigma_{\mathbf{n}}}$$

Where

s = number in any one individual subcategory of response to questions.

$\Sigma_{\mathbf{s}}$ = total of all subcategory responses to each questions.

n = number in any one individual category of personal characteristic.

$\Sigma_{\mathbf{n}}$ = total number of conferences attendees (122).

The appendix contains all the factors so calculated. These factors were then graphed in order to highlight any marked variations. A selection of significant graphs is provided in the following sections.

Motivation for attending conference

In the survey instrument conference attendees were asked why they had enrolled in the course. The open wording of the question invited responses without specific prompts; as a result the responses were wide-ranging. Nevertheless, the responses were able to be grouped into seven categories which were:

- an expressed desire for speaking Bable [40.1% of responses];
- a wish to learn how to write in the language [6.4%];
- cultural motivations [11.2%];
- linguistic activism (ie wishing to defend or promote Bable) [13.9%];
- an interest in teaching the language [18.2%];
- other study or work interests [4.3%]; and
- an intent to overcome a situation of diglossia with respect to the respondent's use/knowledge of the language [5.9%].

The most frequently expressed reason was a desire to speak in Bable, followed by an interest in teaching the language. Linguistic activism was somewhat higher than cultural motivations.

The graph for motivation by age group is indicated in Figure 1. This graph reveals that there was a concurrence of opinion between those in their twenties and those in their forties as to their relative disinterest in cultural motivations for attending compared with the above average interest in this motivation category of those in their thirties.

Indicative of the higher average interest of those in their thirties on cultural motivations was the response of one attendee who stated *"It is sentimental or intimate; it is a criterion of self-identification and a nexus of the conjunction with others"*.[t]

[t] *Ye sentimental o intima; ye ama mena d'autoidentificación y un nexu de xuntura pa colos demas.*

However, those in their forties parted company with their younger fellow attendees when it came to defence or promotion of the language as a motivation, an area where those in their twenties and thirties expressed the greatest interest. One such respondent from this latter group commented *"(it is) to push for that which is our right".*[t]

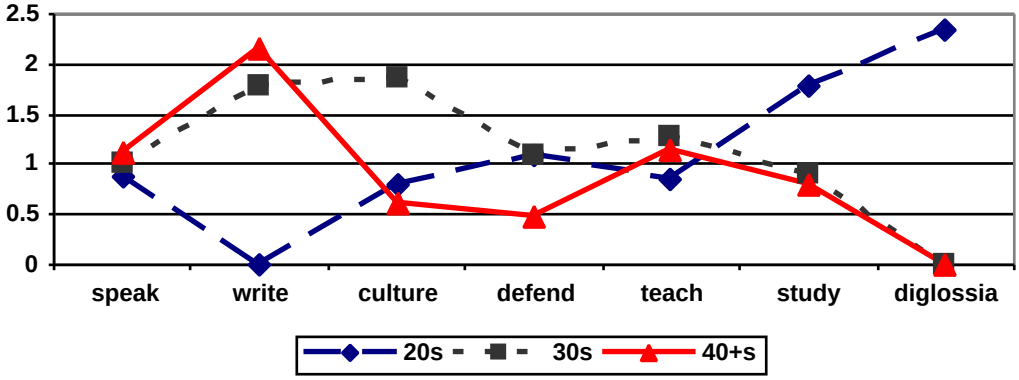


Figure 1: Motivation by age category

The issue of diglossia was noted as an area of concern for those in their twenties. However, this did not mean that older age groups did not believe their use of Bable was diglossic, rather that it had not been a motivation for them attending the conference. They were more likely to be concerned, however, at the quality of the Bable they spoke as a result of interference from Castilian. In response to that category, the following quotes from those in their forties appeared:

> *I don't speak it too well because of the many interferences from Castilian which there are.* [t]
>
> *I believe that I don't speak it 'well' (because it is a very Castilianised Asturian), but the truth is that it embarrasses me to speak it any other way; people don't understand you in Castilian nor in Asturian.* [t]

Motivation according to responses categorised by language zone is mapped in Figure 2. The small sample size for those from the Oriental region (6) and Occidental (10) qualify the conclusions which can be drawn from this graph.

Those from the Occidental zone expressed no interest in defence or promotion of the language as a motivation; whilst those from the Oriental zone took a similar position with respect to cultural motivations (which was contrariwise strongly supported by those from the Occidental zone). Not unsurprisingly diglossia and the related theme of interference were most significant motivationally for those from the Central zone (where the urban context is strongest).

Language ability

More than any other question in the survey, this one was open to subjective interpretation. Respondents were not asked to provide any quantitative measures of language ability. This subjectivity was not seen as a limiting factor in interpreting the responses; rather it was seen as enhancing the capacity to probe attitudes as well as facts. However, some manipulation of the responses was required in order to fit them into a discrete number of response categories.

---

[t] *Puxa polo que nuestru drechu.*

[t] *Nun lo falo bien del too porque son munches les interferencies del castellanu que hai.*

[t] *Creo que nun lo falo 'bien' (porque ye un Asturianu mui Castellanizáu) pero la verdá ye que me da vergoña falalu d'otra manera: si la xente nun ti entiende en Castellan nin en Asturianu.*

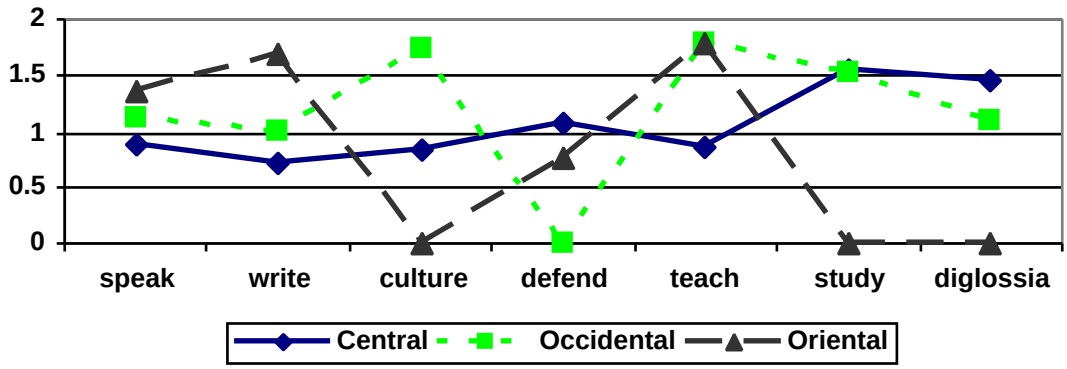


Figure 2: Motivation by language zone

The categories used for grouping were 'Nil' (1.6% of responses), 'Poor' (15.6%), 'Average' (49.2%), 'Good' (17.2%) and 'Very Good' (13.1%).

Gender provided an interesting variation with respect to perceived language ability as shown in Figure 3. With the exception of those who could be included in the category of having self-assessed their language ability as 'very good' (where females gave a higher response rate than males), females generally were more negative about their language ability than their male counterparts. Males were significantly above average in the 'good' category. Given the subjective nature of the responses it is more probable that the differences in the quality of Bable spoken was as much perceived as measurable.

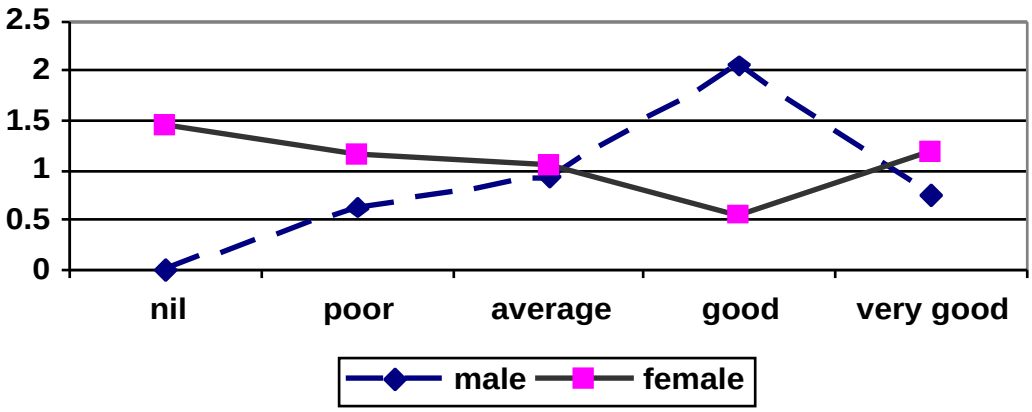


Figure 3: Language capacity by gender

By way of example of different self-perceptions were the following comments from two men and two women respondents. Firstly the men: *"I believe that I speak it well, but not in the normalised form."* [t] and *"I could do it better, but I have made a very large leap (in knowledge of Bable) in recent years."* [t]

Secondly two from amongst the female respondents: *"It seems to me that I don't speak it too badly, though still I could improve it."* [t] and *"I endeavour to use (Bable) whenever I can,*

[t] *Yo creo que lo falo bien, pero non el normaticu.*
[t] *Podría facelo meyor, pero dí un salto pergrande nos últimos años.*
[t] *Paezme que nún lo fal del to mal, anque tovía pueo meyorar.*

above all at home, with acquaintances, in school. No, it does not seem to me that I speak it well." [t]

The analysis of perceived language ability on the basis of language zone was relatively inconclusive; however, the same did not apply to analysis by significant locality of experience (either birthplace, childhood or adulthood location). Figure 4 plots self-assessment according to whether the respondents listed an urban, mining or rural locale as significant in their responses.

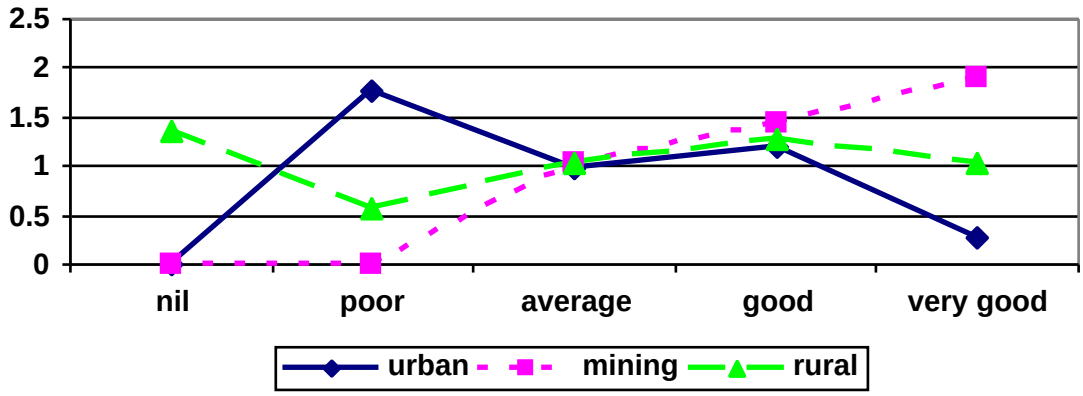


Figure 4: Language capacity by locale

Whilst those with rural connections roughly matched the whole group average, those from mining areas were clearly more positive in their self-assessment of language ability with those from urban areas less positive. It is notable that the mining zones has provided more support for the positive promotion of Bable than rural areas where there has been a tendency for it to be considered a somewhat archaic aspect of their heritage.

Source of language learning

Survey respondents were asked where they had learnt such knowledge of Bable as they had. They were invited to nominate more than one source of learning; however no possible sources were suggested. The responses were then grouped in four broad categories, with one of them being subdivided into two sub-categories ('Mother tongue' and 'Home environment' – 38.8% of all responses). The other three were 'Community' (21.1%), 'School' (3%) and 'Other' (37.1%). This last category included courses offered by local councils (often referred to with the diminutive *cursinos*), university studies and through the reading of books. As indicated earlier, multiple answers were invited. The 122 survey respondents gave 232 sources of learning. Figure 5 plots the spread between learning sources by age grouping.

The most notable variation according to age group was in the 'school' category. The nil response rate to this category from those in their forties and above was the clearly the result of Bable classes only having been offered in Asturian schools since 1984. What was perhaps a little surprising was that those in their twenties gave a lower response to this category than those in their thirties. Further analysis of the responses from those in their thirties who indicated school as being a source of their knowledge of Bable, revealed that all of them were in their early thirties.

While those in their twenties were comparatively less likely to have nominated schools as a source of learning compared with those in their thirties, they were slightly more predisposed to nominate other sources than the latter group.

[t] *Procuro facelo, siempre que puedo. Sobre too n casa, con conocíos, na escuela. Non, nun me parez que lo falo bien.*

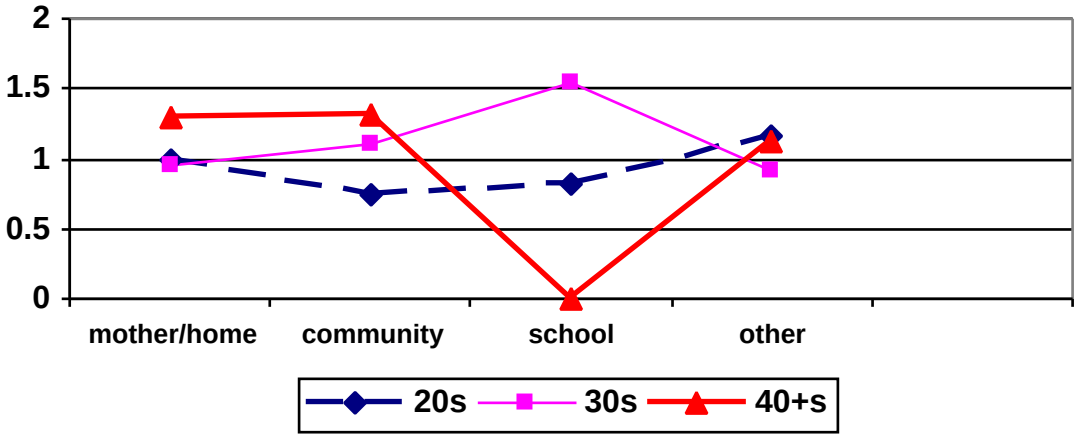


Figure 5: Source of learning by age group

Figure 6 plots sources of learning by locale. The most notable variation which was revealed by analysing source of language learning on the basis of locale was the comparatively higher rate of response for 'school' being a source for such learning amongst those from mining communities than other areas.

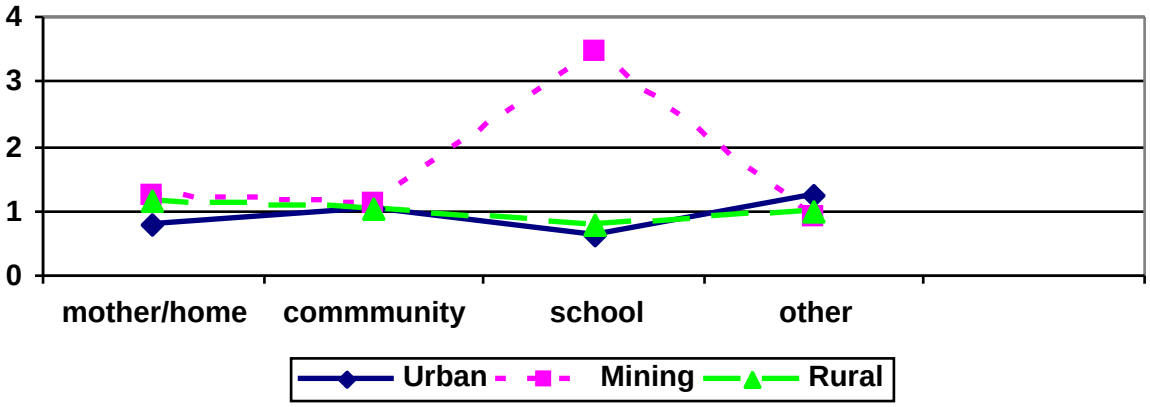


Figure 6: Sources of learning by locale

An interesting feature was a lower propensity for urban-connected respondents to indicate 'mother tongue/home' as a source for the learning of Bable. The average ranking for 'community' and the above average ranking for 'other' amongst this group suggests the relatively greater influence on their learning of Bable of peers and courses offered by local authorities than the home environment. The implication may be that high usage of Bable in the home was not a normal feature in the urban zone, and therefore the language learning of these respondents was by way of a recuperative effort.

Domains of usage

Surveys of Bable usage have always contained generalised statistics as to the levels of usage of Bable; what has not always been done has been an analysis of domains where such usage predominates. The data from this survey was aggregated in four categories: family (37.2% of all responses); 'friends' (34.2%); work (14.1%); and general (14.5%). As multiple responses were common in this group, the total was nearly double the number of attendees (234). Figure 7 plots usage by gender.

Domains of language use analysed by gender revealed that males were more likely to use Bable amongst family and friends than females.

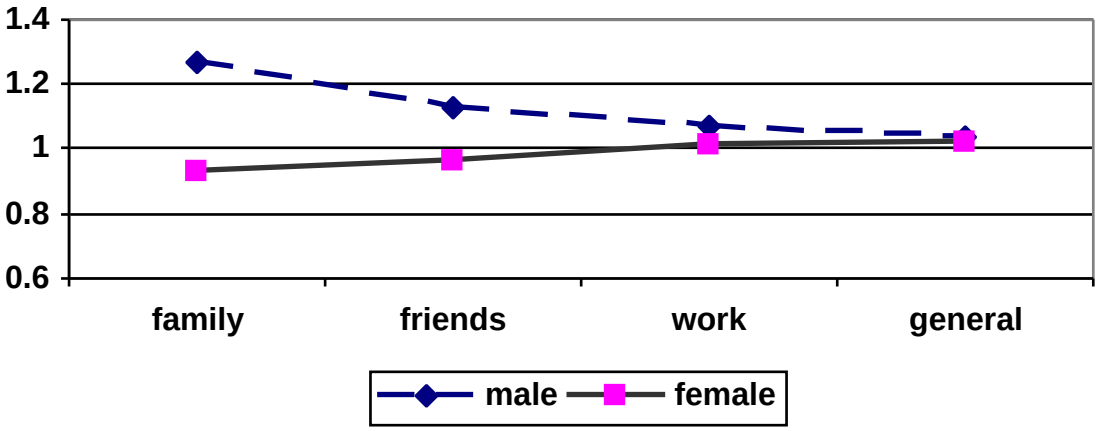


Figure 7: Usage domains by gender

A particularly interesting comment in terms of differential usage of language according to domain was made by one respondent:

> *I speak Asturian only in situations of 'the ghetto', that is, with people who also know the Asturian speech. In other social situations … affectively such as the administrative I use a mixed Asturian. In situations of written communication I utilise Asturian most of the time, some 80% of the situations.*[t]

Figure 8 plots domain usage by age category and reveals that breadth of usage appears to decline amongst younger age groups. Surveys have shown that whilst there has been some increase in the percentage of the population claiming to understand Bable, the increasing depopulation of rural areas and the decline of the mining sector has undermined two key areas where Bable still has any sort of *lingua franca* status. There has been a very significant movement of younger people away from these areas in search of work and lured by the 'streets of gold' of the cities.

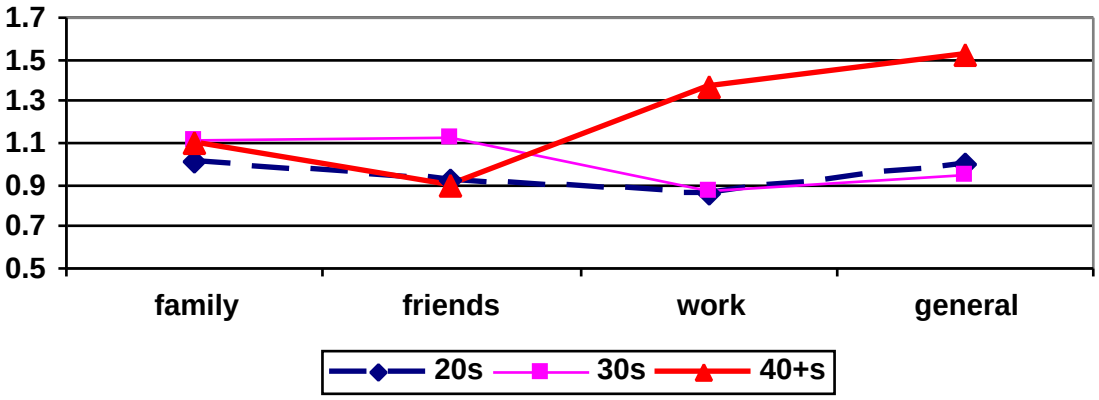


Figure 8: Usage domains by age

Figure 9 plots the usage domains by locale and reveals a clear hierarchy of usage with the greatest relative incidence being amongst those living or previously connected with mining communities; the second highest relative incidence being amongst those connected with the

[t] *Falo asturiano solo nes situaciones 'de ghetto', esto ye, con la xente que tamién 'sé' que fala asturianu. Nel restu de les situaciones sociales … afectivu comu alministrativu usu un asturianu amestáu. Nes situaciones de comunicación escrita utilizo mayoritoriamente l'asturianu, nun porcentax del 80% les situaciones.*

rural environment. Not unsurprisingly urban connected respondents had the lowest relative incidence in the work and general domains.

Exceptions to the norm of relatively lower incidence in the urban locale were invariably from the city of Xixón/Gijón, typified by the following comment made about one respondent's parents: *"My parents are from Xixón and are Asturian speakers".*[t]

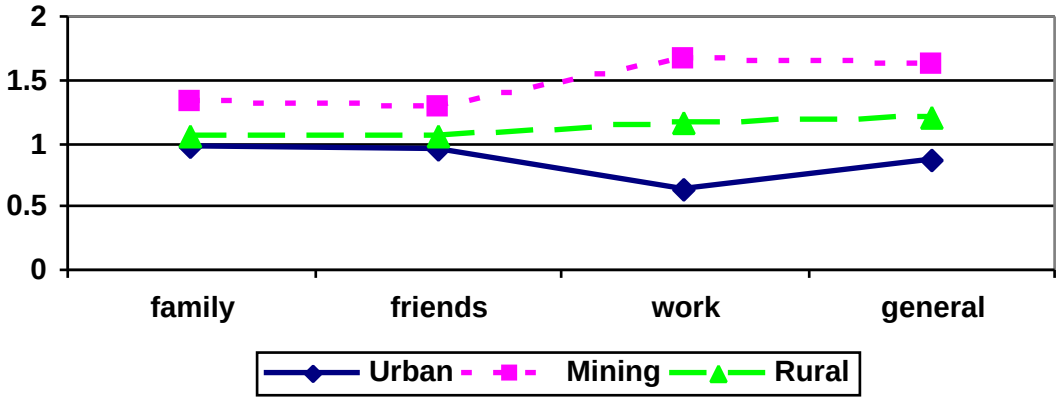


Figure 9: Usage domains by locale

Two comments from mining communities indicated contrasting effects on language use of mixed parentage. In both instances one parent was from Asturias and the other from beyond the region. In the first, the father was from Asturias, in the second, the mother: *"I was born in a mining zone, daughter of and Asturian father and Castilian mother and I am very happy because in my house Asturian is always spoken".*[t] and *"I live in mining zones … my father is from Castille and my mother Asturian, in my house Asturian is not spoken".*[t]

One of the respondents who lived in a rural area made a significant comment of differential usage by domain: *"I speak in Castilian or in Asturian … depending on if I work in the village or in the city".*[t]

QUALITATIVE ANALYSIS

The survey responses also provided a wealth of information when analysed qualitatively. Particular attention was paid to the question of normalisation and officiality as well as to the invitation for respondents to provide a list of ten Bable words. However, in addition some analysis was done of the way in which other responses had been written.

Opinions on normalisation and officiality

120 of the conference attendees responded to both questions on normalisation and co-officiality. One of the remaining two responded just on co-officiality (with the comment *¡Oficialidá ya!* [Officiality Now!]); while the other left the segment blank.

All those who responded to these questions were in support of both concepts. Words like *perimportante* [very good] and *pernecessariu* [very necessary] recurred many times.

---

[t] *Los míos pas son Xixoneses y asturiano-falantes.*

[t] *Nací na cuenca minera, fía de pá Asturianu y ma Castellana y toi percontenta porque na mío casa falen toos l'asturiano.*

[t] *Vivo en cuenques mineres … mío pá ye de Castilella y mío y mío ma asturiana, na mío ca nun se fala asturianu.*

[t] *Falo en castellan o n'asturianu … según trabaye nun pueblu o ciudad..*

In the case of normalisation there were, however, some nuances in the range of supportive comments. One wrote of the significance of this as a *"seña de nuesa identidá"* [sign of our identity] clearly taking a position in the Bable vs Bables polarity. Yet others, still supportive of normalisation, were concerned for the survival of diverse varieties of Bable with such comments as:

> *It is necessary for good communication … (that) the differences between Castilian, normalised Asturian and the distinct varieties be known.*[t]
>
> *I understand that the normalisation of Asturian is necessary from the point of view that there should exist a grammar and some norms of writing. But on the other hand it seems to me that the dialectical variants, rather than prejudicing the llingua, what they do is enrich it.*[t]
>
> *It is necessary if we plan for an institutionalised speech, but it is very important that the local versions are preserved.*[t]

Free word choice

The conference attendees were asked to nominate ten Bable words which occurred to them at the moment of completing the form. The 1220 choices they made fell into 565 word groupings. A word grouping consisted of nouns, verbs and participles of the same base word as well as different spellings. About ten percent (56) of these word groupings were chosen a combined total of 451 times or 37.0% of the total, each with a minimum of five citations.

There were two most frequently cited word groupings, both with 21 citations. The first was *falar/fala/falable* [to speak/speech/'speechness']. Given the theme of the conference, it was hardly surprising that this grouping was the most cited. However, that such citations represented only 1.7% of all words cited perhaps was somewhat lower than might have been expected. Indeed, considering that all language-related word groupings only gave a 4% figure was unexpected. The other groupings were:

> *Llingua* (11 citations)
> *Cooficialidá/Cooficialidad* (7)
> *Asturianu* (5)
> *Normalización/normativización* (5)

The second of the top groupings was *prestar/prestame/prestosu(o)(a)* [to like/I like/likeable]. As with a number of Bable words, *Prestar* is an example of a word whose meaning is different from its homonym in Castilian (where it means *to lend*). It is possible the tension of difference/similarity of this word with Castilian intrigued the conference attendees sufficiently to prompt its inclusion in the responses.

I had expected a reasonably high occurrence of words connected with characteristics of the region – such as the rural, festive, mythological or related to fauna and flora. Some of the words in these and categories related to people were:

> **Flora/fauna:** *Curuxa [owl], Andarina [swallow], Xatu/xatín [calf], Llobu [wolf], Arbeyu [pea], Fueya [leaf], Castañéu [chestnut], Zrezal [cherry tree], Coríu [duck],*

---

[t] *Ye necesaría pa una bona comunicación (que) … se sepa dixebres ente al castellan, l'asturianu normalizáu y les distintes variantes.*

[t] *Entiendo que la normalización del asturianu ye necesaria dende'l puntu de vista de que tien qu'esistia una gramática y unes normes d'escritura. Pero por otru lláu paezme que les variantes dialeutales, ñolle de peerxudicar la llingua, lo que faen, ye enriquecela.*

[t] *Necesaria si pretendemos una fala institucionalizá, pero perimportante que se caltengan les dixebres fasteres.*

Esperteyu [bat], Caparina [butterfly], Carbayu [oak], Escarabaya [beetle], Lloréu [laurel].

Rural: *Horreu [typical barn], Cachu [manure], Gueta [chestnut harvest], Estropayar [to spoil the harvest], Enrestrar [Make a corn string], Abeyera [beehive], Espantaxu [scarecrow*].

Weather: *Abellugar* [*shelter from rain*], *Seronda* [*autumn*], *Orbayu* [*drizzle*], *Borrina* [*mist*], *Nublina* [*fog*], *Xarriar* [*rain heavily*], *Xorociar* [*to rain*].

Mythology: *Trasgu [an elf-like creature], Xana [a mythical female figure], Güestia [procession of the dead]*

Festive: *Folixa [good time], Gayola [good time], Antroxu [Mardi Gras], Amagüestu [chestnut festival] .*

Iconic features: *Madreñas [clogs], Sidra [cider], Gaita [bagpipes]*

Relatives: *Neñu [child], Güelos [grandparents], Pas [parents], Má [mother].*

Other people: *Collaciu [friend], Guahe [lad], Xente [people], Escolinos [students], Maestru [teacher], Muyer [woman], Mozu [youth], Rapaz [lad].*

Further to the example of *prestar* where a word may mean something different in Bable from its meaning in Castilian, a total of six such words were contained in the word lists. They were:

Prestar – to like in Bable; *to lend* in Castilian
Escaecer – to forget in Bable; *to decay* in Castilian (though Medieval Castilian also listed *to forget* as a meaning.
Gayola – good time or *joy* in Bable; *prison* in Castilian.
Seronda - Autumn in Bable; *late fruit* in Castilian.
Andarina – swallow (bird) in Bable; *a female walker* in common Castilian usage,though a *swallow* in rare and archaic usage.
Cai - street in Bable; *wharf* in Castilian (derived from the French *quay*).

Orthographic variation

An interesting feature of the word choices, given the overwhelming support of the group for normalisation of Bable, was the variety of spellings given for some of the words – which was clearly a function of different varieties to which each respondent had been exposed. The following were examples:

Abellugar, aberyugar and *abeyugar* [*to shelter from rain*]
Andarina, andolina, andarines, andorines, andolines [*swallow*]
Galloleru, gayolero and *gayoleru*
Guahe, guaje and *guahe* [*lad*]

Horreu, horru and *horro* [name of characteristic Asturian *barn*]
Maestru, maiestro and *mayestru* [*teacher*]
Neñu(a) and *ñenu* [*child*]

Not included in the responses to the request for ten words, but incorporated elsewhere in their responses were some other diverse spellings such as *mallestra* for *maestra* and *muller* for *muyer* [*woman*].

Use of diminutives

Asturians are known amongst other peoples of Spain for their love of diminutives. The use is almost always in endearing rather than deprecatory senses. Though the survey did not provide for lengthy responses to each question which might, in their length, have been more conducive to the use of diminutives, it was noted that there was usage in the case of such words as *tierrina* [*little land* – referring to Asturias, as opposed to the *big land* which would be Spain]

and *pueblín* [*little village* – referring to the person's own village as opposed to *pueblo* which would be anyone else's]. There was also an array of diminutive forms of the Bable word for *little.* In that latter instance the forms used again showed great variation: *pequinica, pequeñin, pequetín, piquitín, poquitín, pocoñin, poquiñin* and *puquiñin.*

Use of first person plural possessive pronoun

The use of the first person plural possessive pronoun is more frequent in Bable than in English, and reflects a communal spirit traced back to Celtic times. In the case of survey responses such extensive usage was not observed. Significant instances focussed on language (9 cases referring to *llingua* and one to *fala*), culture (5), country (including 2 for *país* and one for *tierra*), identity (2) people (1) and human rights (also 1). Amongst the usages of this pronoun were the following examples: *"The culture of Asturias, it is something ours"*[t] and *"So as not to lose (the language) which is ours".*[t]

There were seventeen uses of the first person singular possessive; eight related to maternal language, eight simply to language and one to land.

Again an interesting feature of the plural possessive was the variety of forms used – namely the forms common in each dialect zone of the region: Bable Occidental (*nueso/a/u*), Bable Central and Bable Oriental (*nuestru/a/o*) as well as that found in both those latter zones and in Castilian (*nuestro/a*).

Usage of Bable & Castilian forms of *to be*

There is a phenomenon in Bable usage patterns referred to as *yeismo.* This the practice whereby a speaker would, in otherwise normal Castilian speech, use the Bable *ye* instead of Castilian *es* for *it is*; in this sense *ye* is something of a sociolinguistic icon in Asturias. One respondent noted such usage: *"at home a mixture of Asturian and Castilian was always spoken (perhaps a little more of the first > the use of pronouns, the 'ye', some verbs)".*[t]

Yeismo represents one of the few examples of common interference of Bable upon Castilian. More commonly interference is in the other direction. Such is also the case with the two forms of the verb *to be* in both Bable and Castilian. Table 1 reveals the incidence of different forms of the two forms of the verb *to be* which appeared anywhere in the survey response forms. It reveals some evidence of Castilian interference, though probably less than would be the case in everyday speech patterns as the respondents may have wanted to avoid obvious interference. Certainly lexical frequency studies such as *Diccionariu de Frecuencies Léxiques del Asturianu*[1] suggest a higher level of such interference than in the table.

One respondent altered a Castilian usage to a Bable one (*soy* > *soi*) while two used the Castilian form of the first person in one verb usage (*soy*) and the Bable in the first person of the other (*toi*).

Usage of one typical Bable form

The use of typical words or forms which can be considered as markers of linguistic identity was frequently seen in the responses from the conference attendees. A number of these were observed in the survey responses.

One example was the use of the prefix *per-* [in the sense of *very* - where *muy* would have been used in Castilian or even sometimes *mui* in Bable]. Examples of such usage were (one instance each unless otherwise indicated): *perbueno, perbaxu, perclaro, percomplicada,*

---

[t] *la cultura d'Asturies, ye algo nuestro.*

[t] *pa no perder lo que ye nuestro*

[t] *En casa siempre se faló un amestáu de asturianu y castellán (quizas un pocu más de lo primero > usu de los pronombres, el ye, dellos verbos)*

[1] Fernando Cuetos *et.al*, *Diccionariu de Frecuencies Léxiques del Asturianu*, Academia de la Llingua Asturiana, Uviéu, 1997.

percontenta, perdiferente, perdificil (2), *perfechu (2), pergrande* (2), *perguapa* (2), *perimportante* (28) and *pernecesaria* (13), *perprobe, pertriste, perverde.*

Table 1: Comparison of some Bable & Castilian forms of verbs 'to be'

Bable		**Castilian**	
*Ser (*17) [to be (self)]			
Soi [I am]	86	*Soy*	14
Ye [S/he/it is]	244	*Es*	4
		Somos [We are]	2
		Son [They are]	33
		Fui [I was]	2
Foi [S/he was]	13	*Fue*	1
		Fuimos [We were]	1
Yera[I was (imperfect)]	8	*Era*	1
Yeren [They were]	5	*Eran*	0
Sedre [I will be]	1	*Sera*	0
Sedra[S/he will be]	2	*Sera*	0
		Seria [I, he/she/it would be (conditional)]	2
Seya[I w'd be(subjunctive)	0	*Sea*	1
[form doesn't occur]		*Fuese* [I were (subjunctive)]	1
		Fuera [I, he/she/it were (subjunctive)]	2
Tar (11) [to be (place)]		*Estar* (0)	
Toi [I am (place)]	87	*Estoy*	0
Ta [He/she/it is]	18	*Esta*	0
Tamos [We are]	6	*Estamos*	0
Tan [They are]	4	*Están*	0
Tea [I would be]	3	*Este*	0
Taba [I, s/he was]	1	*Estaba*	0

[Notes: Where Bable and Castilian share the same form, the form has been centred across both columns.]

Per- is not always used in Bable, in many cases *mui* would be more appropriate; Milio Cueto, in his work *Manual de la Llingua Asturiana* noted that such overuse of the *per-* form in place of the word *mui* represented hypercorrection on the part of Bable speakers trying to compensate for interference pressures from Castilian.

In answers to the general questions there were four instances of the more traditional Castilian forms (though still using the Bable variants in both cases) of *mui bien*, *mui importante* and *mui necesaria*. In one instance, a respondent used two different types of mixed Bable and Castilian forms when he wrote *perbien* in the context of normalisation and *mui bono* in the context of coofficiality.

Rural affiliations and other sentiments expressed

A number of survey respondents made comments indicating a rural sentimentality. For example: *"I live in Uviéu though I feel myself affectively linked with the rural zone.*[t]*"*

[t] *Vivo n'Uviéu magar siéntome afeutivamente lligáu cola faster a rural.*

Two respondents noted feeling fear or embarrassment in certain contexts: *"Fear at being ridiculed on the part of others"* (in clarification of comment that he did not use Bable as often as he would like).[t]

One noted the negative treatment accorded to those speaking Bable in the school system in the 60s and 70s: *"Today I would speak it totally if it had not been that I kept it hidden at school in the 60s and 70s where they didn't respect it"*.[t]

One noted the fiesta-connection which often applies to the use of Bable: *"(I speak it) when we are all together as a family in some family fiesta"*.[t]

CONCLUSION

The survey of a group who might be considered partisans for the recuperation of Bable as a sociolinguistic force revealed both varying motivations and usage patterns. It can be assumed that these variations would be no less marked in the general population.

The various motivations, though often consonant, represented a range of opinions which reflected the varying circumstances of the participants; and which could not easily be reduced to simple, homogeneous statements.

The usage patterns reflected the fact that at this stage Bable is still for many practical considerations in a pre-normalised state. They also reflect the hazards of the close linguistic relationship existing between Bable (as an L1) with Castilian (as an L2).

Policy decisions which focus on normalisation and officiality need, therefore, in the case of Bable to reflect these local characteristics rather than be replications of policies tried elsewhere where motivation and usage patterns will have been different.

REFERENCES

Cueto, M.R. (1994) *Manual de la Llingua Asturiana,* Llibros del Pexe, Xixón.

Cuetos, F. et.al (1997) *Diccionariu de Frecuencies Léxiques del Asturianu,* Academia de la Llingua Asturiana, Uviéu.

[t] *de ciertu miéu a facer ridículo per parte la otra xente.*

[t] *Güei falariala dafechu si nun me lo fixera escaecer el colexu nos años 60 y 70 onde nun lo respetaben.*

[t] *cuando tamos reuniós toa la familia en dalguna fiesta familiar.*

APPENDIX

Personal characteristics	n =	Motivation for attending							Language capacity					Where learnt					Usage			
		Like to speak	Learn to write	Cultural	Defend/promote language	Want to teach	Further study/work	Overcome diglossia	Nil	Poor	Average	Good	Very good	Mother tongue	Home environment	Local community	School	Other	Family	Friends/c'ty	Work	General
s=	122	75	12	21	26	34	8	11	2	19	60	21	16	20	70	49	7	86	87	80	33	34
Gender ns	7	0.93	1.46	-	2.7	-	-	-	-	0.91	0.88	1.67	-	2.63	0.51	0.72	-	0.4	0.6	0.88	0.53	0.51
Male	31	0.89	1.31	1.13	1.21	0.46	2.95	1.07	-	0.62	0.92	2.06	0.74	1.18	1.18	1.13	1.69	1.05	1.27	1.13	1.07	1.04
Female	84	1.04	0.85	1.04	0.78	1.28	0.36	1.06	1.45	1.15	1.04	0.55	1.18	0.8	0.97	0.98	0.83	1.01	0.93	0.96	1.01	1.02
Occup. ns	28	1.16	1.45	0.83	0.83	0.64	1.09	1.58	2.17	1.37	0.8	1.03	0.82	1.09	0.74	0.71	2.48	1.06	0.75	0.65	0.4	0.77
Other	11	0.89	0.92	1.06	1.71	0.32	-	3.03	-	1.17	1.11	1.06	-	0.56	1.43	1.13	-	1.29	1.02	0.83	1.34	0.66
Teacher	66	0.96	0.92	1.14	0.92	1.47	0.46	0.34	0.92	0.97	1.05	0.88	1.27	0.92	0.92	1.06	0.53	0.97	1.08	1.11	1.18	1.03
Student	17	0.96	0.6	0.68	1.11	0.21	3.6	1.31	-	0.38	1.08	1.37	0.9	1.44	1.23	1.17	1.03	0.83	0.91	0.99	0.87	0.85
Age ns	26	1.13	0.78	0.45	1.08	0.83	-	0.85	-	0.74	0.78	1.34	1.47	1.17	0.8	1.05	1.34	0.77	0.81	1.06	1.14	0.69
Twenties	43	0.87	-	0.81	1.09	0.84	1.78	2.33	1.42	1.2	0.99	1.08	0.71	0.99	0.97	0.75	0.81	1.16	1.01	0.92	0.86	1
Thirties	34	1	1.79	1.87	1.1	1.27	0.9	-	-	1.12	1.14	0.68	0.9	0.72	1.03	1.1	1.54	0.91	1.11	1.12	0.87	0.95
Forties+	19	1.12	2.15	0.62	0.49	1.14	0.81	-	-	0.68	1.08	0.92	1.21	1.29	1.28	1.32	-	1.12	1.1	0.89	1.37	1.52
Zone s/ext	36	1.08	1.41	1.29	1.04	0.9	-	0.31	3.39	1.61	0.74	0.64	1.06	1.36	0.73	0.55	-	1.06	0.78	0.81	1.03	1
Central	69	0.9	0.74	0.84	1.09	0.88	1.55	1.45	-	0.93	1.09	1.26	0.77	0.71	1.06	1.15	1.77	1.03	1.10	1.1	1.07	1.04
Occidental	10	1.13	1.01	1.74	-	1.79	1.52	1.11	-	-	1.43	-	1.52	-	1.57	1.49	-	0.85	0.98	0.91	0.74	0.72
Oriental	6	1.37	1.69	-	0.78	1.80	-	-	-	-	1.02	1.94	1.29	4.08	0.88	0.84	-	0.71	1.16	1.02	0.61	1.2
Ast-Gall	1	1.63	-	-	4.75	-	-	-	-	-	-	-	7.88	-	1.75	2.5	-	-	1.38	1.63	-	-
Ext/ns	28	1.16	0.73	0.83	1.04	0.9	0.54	-	2.17	1.6	0.94	-	1.09	1.30	0.81	0.71	-	0.81	0.7	0.76	0.66	0.38
Urban	29	1.01	0.7	0.8	0.84	0.87	0.53	2.67	-	1.77	0.98	1.2	0.26	0.63	0.84	1.03	0.6	1.22	0.97	0.95	0.64	0.87
Mining	20	0.90	1.02	1.16	1.46	0.54	3.05	0.55	-	-	1.02	1.45	1.91	0.91	1.31	1.24	3.48	0.92	1.33	1.29	1.66	1.62
Rural	45	0.94	1.36	1.16	0.87	1.36	0.68	0.74	1.36	0.57	1.04	1.29	1.02	1.08	1.08	1.05	0.78	1.01	1.06	1.05	1.15	1.2

Did they get what they came for? Evaluating teachers' learning

Kerry Bissaker
Flinders University, School of Education *kerry.bissaker@flinders.edu.au*

Twenty-five teachers enrolled in a fully funded Graduate Certificate program in the area of Learning Difficulties were asked to comment on their motivation behind applying for the course and what they hoped the outcomes of their study would be. Research on quality teacher professional development, along with consideration of the participants' responses, were used to inform the design, aims and content of the course. Outcomes were evaluated using Guskey's (1999) model and considered participants' reactions, participants' learning, organisational change and support, participants' use of new knowledge and skills, and student learning outcomes. Following the final topic in the course participants were asked to review their pre course responses and reflect on whether they felt their intended outcomes had been achieved. The participants were also encouraged to consider the value of such a course in their development. Responses from the participants clearly highlighted that student learning outcomes had benefited along with development of their own knowledge and skills.

Teachers' learning, professional development, evaluation

INTRODUCTION

Professional development for teachers is clearly recognised as having a significant influence on promoting positive learning outcomes for students. Sykes (1999:152) notes that 'whereas professional development initially was regarded as one among a number of coequal policy instruments for promoting change, it is now reckoned as the centrepiece'. The Australian Federal government has provided significant financial assistance to support the initiative, *Teachers for the 21st Century*, basing its support on the research that 'confirms [the] value of professional development in raising educational standards' (DETYA, 2000). Further evidence of system support for professional development is reflected in South Australia where teachers are provided with an additional week of vacation if they complete 35 hours of professional development throughout the year.

Research on professional development presents a dichotomous view on optimal approaches to professional development, noting that many traditional models of professional development are considered fragmented and poorly coordinated (Guskey, 1995, Cohen and Spillane, 1992 cited in Sykes, 1999). Often little thought has been given to the strategic application of knowledge and skills presented in professional development programs. Many professional development programs are presented as 'one hit wonders' with a focus on the latest 'trend' (Hawley and Valli 1999). If governments are to fund professional development, then providers must design effective programs that can evaluate outcomes for both teachers and students.

Increasingly, it is recognised that there is no one 'perfect' approach to successful professional development because the content, process and contextual variables will differ across programs, styles of delivery and learning, and situation. However, Hawley and Valli (1999:137) describe eight characteristics of effective professional development. Some of these characteristics include:

- teachers clearly identifying their learning needs
- processes that involve collaborative problems solving

- organisation based on the continuous and ongoing involvement of a cohesive group
- opportunities to develop theoretical understanding of new knowledge and skills
- integration of professional development within a comprehensive change process including the facilitation of student learning
- incorporating evaluation of multiple sources of outcomes for teachers, students and organisations.

All of these characteristics provided the foundation for developing the project described below. However, the final aspect, that of evaluation, is the focus of this paper. Guskey (1998:1) suggests that while we engage in evaluation on a daily and ongoing basis, the evaluation of professional development requires a systematic approach that is 'thoughtful, intentional and purposeful'. Employing systematic evaluation, that could inform the provider about quality of content, processes and outcomes, was a key consideration of the project described in this paper, particularly in justifying the financial support provided by the South Australian government.

Elmore and Burney (1999:263) suggest that while we know a great deal about the features of effective professional development, 'we know a good deal less about how to organise successful professional development so as to influence practice in large numbers of schools and classrooms'. Although the project described in this paper would not be considered large scale, evaluation processes could provide feedback about the potential of this model in contrast to other forms of professional development. The evaluation process employed for the project was based on the work of Guskey (1998, 1999) who identifies critical stages and levels of evaluation. Guskey suggests that evaluation must be considered at the planning, formative and summative stages. He also suggests that outcomes of professional development should be evaluated at the following levels:

- Participants' reactions
- Participants' learning
- Organisation support and change
- Participants' use of new knowledge and skills
- Student learning outcomes

The project described below used a variety of processes to evaluate the different stages and levels. Planning evaluation sought to establish a data base of information for assisting in the design of the program and also for reflecting on desired outcomes. At the formative and summative stages information was gathered to evaluate outcomes at all five levels detailed above. Processes for doing this are detailed in sections below and appendix one provides Guskey's (1999) model for ways of gathering this information. The model also provides details of how the information can be used. In this project the information was used to: improve program delivery; ensure content was meeting the needs of the participants; document outcomes for the South Australian Department of Education, Training and Employment (SA DETE); and demonstrate the overall impact the program had on the participants, students and their educational environment.

THE PROJECT

The South Australian Department of Education, Training and Employment (SA DETE) Student and Professional Services section, in response to an enterprise bargaining outcome, sought to establish professional development programs in the area of special education. In collaboration with Flinders University, teachers were provided with the opportunity to complete a Graduate Certificate program in a specialised area. Learning difficulties was the specialisation for the project that forms the focus of this paper. Participants' course fees and some teaching release time were provided by SA DETE.

The Graduate Certificate was divided into three topics completed over a period of 18 months. The first two topics used a lecture and workshop process with the participants meeting together every four weeks for either one or two day intensive sessions. Participants completed activities in their own educational settings between sessions and opportunities for sharing feedback was provided in ongoing sessions. The final topic engaged the participants in an individually designed action research project. During this topic students met once at the commencement of the topic. Ongoing support was provided by facilitators who checked on progress, arranged for small groups to meet and provided telephone and email contact. The participants met again at the end of the topic when they presented the outcomes of their research to each other.

PLANNING EVALUATION

The design of the course provided the initial evaluation phase. As Guskey (1998) suggests program design is critical in defining a 'precise understanding of what is to be accomplished, what procedures will be used, and how success will be determined'. This 'planning' evaluation was paramount considering the collaborative process that had been established between SA DETE and Flinders University. A program committee was established to specify the goals of the program, content focus, resources available, time frames and processes of delivery and evaluation. This committee included University and SA DETE personnel, parents of students with learning difficulties, classroom teachers and community consultants in the area of learning difficulties. Each topic engaged a different committee.

The design process allowed the course coordinator of the program to identify SA DETE, community and parental perspectives about critical content. This macro perspective provided a sound basis for developing content goals. However, research clearly indicates that professional development must also consider the micro perspective (Harris, 2000, Kemmis & McTaggart, 1988). The needs of the participants demanded consideration if successful outcomes were to be achieved. To assist in the 'planning evaluation' phase all participants were asked to provide a statement about their motivation in applying for the course and to identify what they hoped to gain from participating in the program. These statements provided additional information to assist in the design of the topics. The statements were also used in a post evaluation process where participants were able to reflect on their initial statement and respond to whether they felt the program had addressed their individual goals. Details of these outcomes will be provided later in this paper.

Participants' responses as to their motivation in applying for the course included the following:

- *"At times I have been unsure of what I could do to further assist students with learning difficulties and have been frustrated with current resources and assistance available"*
- *"There is hardly anyone with specific training in this area (country location) and I feel a person with these skills would be valuable to the school and wider community"*
- *"I am keen to broaden my knowledge with more theoretical understanding and I also wish to further my career opportunities"*

Reponses to outcomes participants hoped to achieve from their involvement in the course included:

- *"I hope to learn more about identifying learning difficulties in students and then to develop effective programs to cater for these difficulties and I would hope to learn many practical strategies that I can use in the classroom"*
- *"Access to current research, practices and resources, and the opportunity to meet other education workers to share ideas and resources"*
- *" I would hope by completing this course I will improve my employment opportunities and update my qualifications"*

All responses from participants could be categorised into three key areas, 1) the development of knowledge, skills and effective practices, 2) the opportunity to develop networks with other participants and 3) to upgrade qualifications and improve employment opportunities. One further outcome that was not generated by participants but by the coordinator of the program, was to assist participants to become reflective practitioners. These desired outcomes along with information from the committees provided the foundation for the content, aims and process of delivery. The information also assisted in designing an ongoing evaluation process to assess whether these aims were being achieved. Guskey (1998) suggests formative evaluation should follow planning evaluation.

FORMATIVE EVALUATION

Formative evaluation took place at multiple stages throughout the program using a variety of written and verbal feedback. The formative evaluation process provided valuable feedback on aspects of delivery enabling the program coordinator to respond with changes in time frames and delivery models. The feedback also allowed for individualising content and consideration of specific contextual issues. Finally the process allowed for monitoring of initial topic goals.

The formative evaluation process included a mixture of open ended and Likert scale questions to rate particular aspects of the course, for example, 'Rate the value of this session in relation to your current educational setting', 1 being extremely valuable with 5 being of little or no value. Open ended questions included, 'Comment on the most valuable aspect of the session from your perspective' and 'What questions remain uppermost in your mind at the end of the session?' Angelo & Cross (1993) suggests that such questions allow the coordinator of programs to respond to issues and aspects that remain of concern for participants.

The feedback from participants could be categorised into the levels provided in Guskey's model. Examples of information from the participants at each level that reflect varying perspectives include the following:

Participants' reactions

- *"A wonderful day. Presentations used humour, related to lives and a sense of practicality. I can easily see how this could fit into my program and teaching"*
- *"Breaking into secondary/primary groups on a number of occasions increased relevance"*
- *"Overload for me on Friday, I'd like to see a bit more interaction between presenters and learners"*
- *"Great lunch and lollies, they kept me going in the afternoon".*

Participants' learning

- *"I now feel I understand what cognition really means"*
- *"It's great to be getting strategies that I can actually implement with children in my class. I was having problems with the retention rate of one of the students and had run out of ideas"*
- *"I actually found myself being able to put some of the different things we've talked about in past sessions in perspective, like the 'penny dropped' and I could see how the different topics are starting to interrelate and make sense in the big picture".*
- *"What are the consensus conclusions on the causes of ADHD and best management practices? I'm still not clear on this"*
- *"How do models of learning match the 8 models of intelligence that I have been hearing people talk about?"*

Organisation support and change

- *"I have used a number of the 'tips for teachers' from the Bos & Vaughn readings to pass out to teachers. I have conducted a workshop on modifying assessment plans for students with learning difficulties and used a number of the ideas from the course. Later this term I am running a session on strategies for students with ADHD as well."*
- *"I've talked a lot to colleagues about the assignments and then this has extended into the why's and wherefore's of the theory. It actually has been quite stimulating discussing what teaching and learning is about rather than what this student did last lesson"*
- *'I feel I have been able to help some staff move from not looking at the learner as having a deficit, changing attitudes to students who don't fit the mould. I have also been able to help teachers choose appropriate testing techniques and how to use diagnostic tests for planning intervention strategies".*

Participants' use of new knowledge and skills – many of the comments detailed in the levels above provide evidence that the participants were using new skills and knowledge. In addition specific details of the effectiveness of the application of new skills and knowledge were also sought. Assignment requirements provided a valuable source for evaluation at this level. Participants were required to complete a case study. Using information gained from appropriate assessment, an intervention program was designed for a specific student in each participants' class. Reflections on implementation of the program and subsequent outcomes were also required in the assessment. Some comments from the participants about this process are recorded below.

- *"I found the opportunity to analyse 'Tom's' learning difficulties in such detail, and then planning for specific outcomes, a valuable process, the assessments found problems that I hadn't specifically noticed"*
- *"Writing specific learning outcomes for the students with LD is a process I will use again as it provided a focus for my teaching and also showed that these kids in my special class could indeed make progress even if they were only small steps".*

Student learning outcomes – evidence at this level was also provided through the assignment requirements. Some participants completed standardised testing to establish benchmarks prior to the implementation of an intervention program. Examples of student learning outcomes included increases in spelling and reading ages, improved scores on mathematics assessments and the use of specific comprehension strategies. Other outcomes that were noted included one student using specific language to gain entry to a group, behaviour contracts being achieved and one class displaying explicitly taught group skills. The most significant examples of student learning were evident in the outcomes of the participants' action research projects. However, these outcomes should be viewed as summative outcomes as there was no opportunity to influence participants' knowledge and practices when projects were being presented. In contrast feedback on assignments could be considered as formative evaluation as the participants could use the feedback to further modify and reflect on their practices prior to commencing the final action research topic.

The formative evaluation process is critical to meeting the participants' needs and the program's objectives. The coordinator of the course was able to respond to many of the participants' suggestions for improving aspects of content and delivery. Guskey (1997) commented that evaluation questions often only seek feedback at Level 1 or 2, rarely seeking to investigate the results of a participant's involvement to the wider school community. The fact that this course was 18 months in length allowed for questions and activities to be designed that provided information at the five different levels. This would not be possible if teachers were only involved in short term (for example, a day long session) professional development activities.

Although many of the comments from participants could be viewed as summative in one sense, the model used in this evaluation process considered all feedback from participants as formative if there was an opportunity to make changes to the program delivery and/or content of the course. Collection of information at the end of each topic, and at the end of the course was considered summative evaluation as there was no further opportunity to respond to the information, except when designing a course for a new group of participants.

SUMMATIVE EVALUATION

Summative evaluations were completed at the end of each of the three topics. This evaluation sought responses to a range of statements including the value of the topic, whether aims were met and participants' satisfaction with scheduling of work requirements. Two open ended questions were also included. Participants rated statements using a Likert scale. A rating of 7 indicated the participants 'strongly agreed' with the statement while a rating of 1 indicated they 'strongly disagreed'. The first two topics delivered involved a lecture and workshop model of delivery and participants responded to the same summative evaluation form. The format of the summative evaluation was based on a model used for evaluating all topics at the University level. However, the standard statements used in this procedure only provided evaluation at the first two levels of Guskey's model.

Responses to the open ended questions provided more valuable information than the responses to the statements, although responses tended to provide evaluation at only the first three levels of Guskey's model. Some of the participants' responses about the best aspects of topic one included:

- *"This course is answering a lot of questions I had before I started. I was feeling really aimless before. Now I feel challenged again".*
- *"I really valued the group work and discussion that came as a result of sharing different ideas".*
- *"I think for me the best aspects have been how directly the study has connected with my work – informed it, developed it and also increased my confidence and ability to present PD sessions to other staff".*

Comments on how the topic could be improved included:

- *"Better airconditioning, more reviewing of what we did in previous sessions".*
- *"I would like greater access to library resources. I feel a little isolated from information. Maybe some time spent in the library would assist me in getting a better picture of what is where. I liked the format of the session even though I normally hate group work".*

Some of the participants' comments in the summative evaluation did not differ significantly from information they provided in formative evaluations. However, the comment above about 'access' to library resources being provided as a summative evaluation represents the difficulty in responding to participants' needs if this information is only provided after the event. The information may prove helpful to planning evaluation for a following program but there is also the aspect that individuals in this next program may have different needs.

The nature of the final topic involved the participants in self-generated action research projects. This allowed for some variation in the summative evaluation process. In particular the coordinator sought feedback from participants at the five different evaluation levels. Eight Likert scale statements and seven open ended questions were included. The open ended questions participants' responded to included the following:

- I feel this topic could be improved by…..
- Describe the most important outcomes of this topic for you personally.

- Describe the most important outcomes for your students based on your involvement in this topic.
- How do you feel your involvement in this topic has contributed to your professional disposition?
- What format of topic delivery did you prefer most, and why? (Lecture/workshop format or Action Research)
- Provide some reflections on the difference between engaging in tertiary study as opposed to engaging in short term professional development.
- List some of the elements that you believe are critical to quality professional development.

The final 3 questions above were specifically designed to seek the participants' opinions about issues beyond the course including the view of quality professional development. The questions were asked to provide information to the course coordinator and DETE about the value of such a model of professional development. It was noted that the participants' identified elements of quality professional development consistent with those suggested by Hawley and Valli, (1999) earlier in this paper. Feedback was also sought on the participants' view of the action research process. Hopkins (1993) supports the notion that teachers engaging in action research develop an increased responsibility for their actions and 'create a more energetic and dynamic environment in which teaching and learning can occur'. Encouraging the participants to develop as reflective practitioners was a major aim of the course coordinator and summative evaluation sought to establish if this had occurred. The focus of this paper restricts further discussion about participants' responses to these questions. However, responses indicated that participants valued their opportunity to engage in action research and commented on developing as reflective practitioners.

The open ended questions within the summative evaluation of all topics provided a rich source of evidence. However, it is interesting to note that most comments made by the participants only reflected outcomes at the evaluation level of 1 and 2. There were some comments that reflected outcomes at other levels. Examples of feedback at the different levels include the following:

Participants' reactions – gained from asking students how the felt the topic could be improved.

- *"Working alone was difficult"*
- *"An extra release day for preparing the presentation"*
- *"Don't ask me, I thought it was all great*"

Participants' learning – feedback was gained from asking the participants to describe the most important outcome of the topic for them. Responses to this question were compared with the participants' pre course desired outcomes.

- *"I now have a successful, manageable and functional program in my classroom which is continually changing and improving as I continue to learn and make connection. This is because I have developed into a reflective practitioner".*
- *"I learnt so much about teaching strategies, particularly the importance of phonological awareness and developing auditory memory for all students".*
- *"I now have a qualification in this area and improved employment prospects".*

Organisation support and change – feedback at this level came from a variety of sources including responses describing the most important outcomes for participants and also from their summative reflections in the action research journals.

- *"My participation in this topic has helped me define and guide the development of my school based position. The outcomes have been more effective use of my time with positive feedback from my colleagues"*
- *" I have increased recognition at the school as someone with relevant skills in the area of LD. The parents also have someone to relate too."*
- *"Through gaining a new position with the LD support team I am able to share my knowledge with teachers across the state".*

Participants' use of knowledge and skills – feedback at this level was generally gained from the action research documentation and presentations.

- *"I have realised that my planning for collaborative learning must be thorough and explicit. Too often I have been setting 'group work' which is more about classroom organisation and children cooperating by sitting in groups, rather than tasks demanding collaborative learning."*

The following comment is a response to outcomes from an action research project involving the use of anti bias language in a pre school setting.

- *"I believe I now understand and feel some confidence in challenging bias. I have seen how bias can restrict children's access to the curriculum. I am also more aware of ways of observing children and their play. I use more modelling now and don't assume children know how to play".*

Student learning outcomes – feedback at this level was generally gained from the request for participants to describe the most important learning outcome for their students but the action research documentation also provided a powerful source of information for this level of evaluation. The first comment, from the anti bias action research project featured above, reflects an important learning outcome at a critical developmental stage. The comment is an observation made by the teacher in her final observation process for the project.

- *"Josh approaches a group of boys who are playing with a ball. He watches, then some discussion takes place and Josh states, "but that's not fair". Josh steps into the middle of the circle and addresses the child holding the ball and states again, "but that's not fair only him getting a go". The child with the ball sighs and states, "okay, everyone can play, but then me and Luke want a turn on our own".*

The following comment was a participant's description of the most important learning outcomes for her students.

- *" Having an informed educator to work with them who is very involved with their best interests and who is deeply committed to their learning".*

Other comments that relate specifically to outcomes for students include:

- *"Increased experiences of success in maths including increased automaticity for 5 out of 7 students with learning difficulties using the Westwood(1995) minute maths tests".*
- *"Increased work output with a more positive attitude towards what they have to do".*

Further aspects of summative evaluation included participants rating on a scale from 1 to 9 their perceived skill and knowledge level pre and post the program. Figure 1 illustrates these ratings. Every participants' post program rating was higher than their pre program rating. The increases in rating varied for individuals with the smallest change being a 1.5 increase and the greatest being 6.5. The participants whose pre ratings were the highest recorded the smallest increase. This may indicate their prior knowledge was a variable to consider in post rating scores. These ratings provide limited information in contrast to that provided by the use of Guskey's model. Indeed such ratings only serve to provide information at level 2, that of participants' learning. However, details about the quality of the learning is absent in such forms of evaluation. A provider using such information may only be able to comment that

there was an increase in participants' knowledge and skills but could not elaborate as to how this increase affected students learning outcomes or organisational support and change.

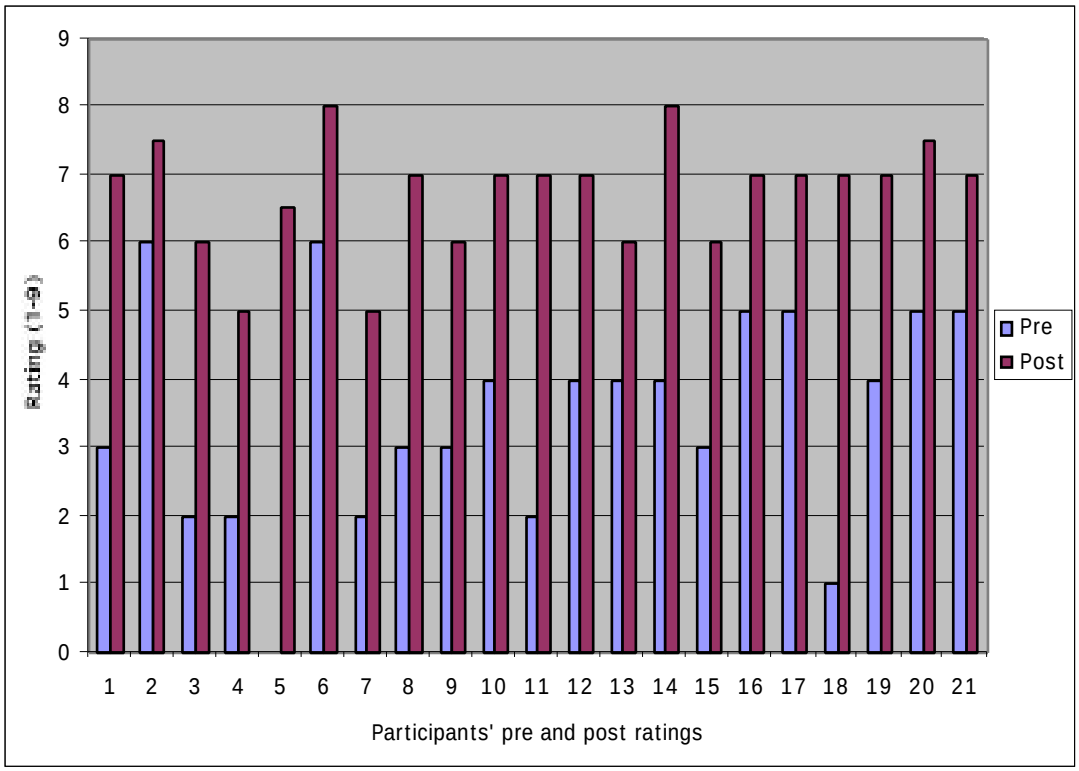


Figure 1: Participants' pre and post program ratings for level of knowledge and skills

Participants were also asked to comment on roles they felt confident in undertaking, having completed the course. Such roles included, providing support for colleagues in the area of learning difficulties, presentations at cluster group meetings, school leadership roles or participation in school management teams and finally, presentations at conferences. Responses to this question provided information to evaluate outcomes at level 3 of Guskey's model, that of organisational support and change. Figure 2 illustrates the number of participants who felt confident about undertaking specific roles following their participation in the course. It is interesting to note that while most participants felt confident about providing support for colleagues fewer participants felt confident about presenting at conferences. Several participants added comments next to their nominations indicating that they had never presented at a conference before and viewed this as frightening prospect for which they did not feel ready. Others felt they would prefer to present at staff meetings and cluster groups prior to considering a conference presentation. These responses need some further investigation by providers and systems, as an additional level in Guskey's model could be to generate information beyond the school level to the wider community. If systems intend to support teacher development then teachers sharing their learning and experiences with other teachers beyond the school environment would appear to be a desired outcome too.

A final element in the summative evaluation process involved the participants reflecting on the comments they made prior to the course and responding to whether they felt their desired outcomes had been achieved. This reflective activity also resulted in some participants identifying previously unrecognised learning opportunities that arose from their participation in the program. The opportunity to reflect on their pre program statements can be noted in the specific details about desired outcomes in some of the following comments:

- *"I wrote in my initial statement that I wanted to develop confidence in my ability to work with students with learning difficulties. My support prior to the course had been very 'ad hoc' and I didn't have a sound knowledge base and understanding of methodologies used*

for students with LD. I now feel very confident in my ability to make a difference for these students".

- *"I stated initially that I specifically wanted to develop motivational techniques for my students in a secondary setting, the most helpful thing I have learnt in respect to this is to give control and responsibility back to the students, it has worked wonders for my students".*
- *"I think overall the course has returned far more than I originally expected, especially in making links with others and in opportunities to write and reflect".*
- *"My desire to broaden my knowledge with more 'theoretical learning' has certainly been achieved. I have been bombarded with a huge amount of valuable and interesting information which motivates me to continue with further studies".*
- *"The course has been supportive of all my desired outcomes and in addition has provided me with an extensive range of resources and practices to use".*
- *"My original outcomes have been achieved as I have read more widely and been updated on current methodology. I think I still need more work in the maths area though this was not stated as an original outcome".*

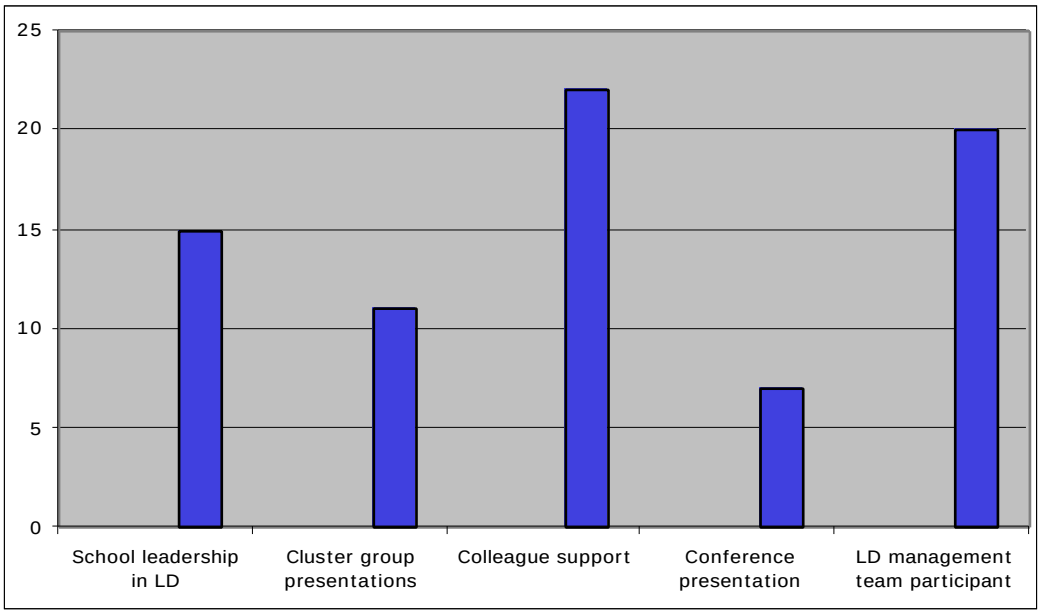


Figure 2: Participants' confidence for undertaking specific roles

DISCUSSION

The participants' stories provided a positive picture from a number of perspectives. Significant outcomes for participants, their students and in many cases, organisations have been achieved from the perspective of the participants. One participant was able to establish a whole school model of tracking students' progress and structuring effective support processes as required, without the need for additional staffing resources. Another participant won a position as a project officer to provide statewide professional development and policy advice in the area of Learning Difficulties. Outcomes for students that could also be noted in the participants' feedback included both individual and whole class improvements. Participants' energy levels to complete the program were often driven by the knowledge that they were making a difference for their students. Guskey (1998:9) notes that *"…anecdotes and testimonials....are typically biased and highly subjective. Nevertheless, they can be powerful and convincing. They are an important source of evidence that should never be ignored"*. He goes on to suggest that *"good evidence is not hard to come by if you know what you are looking for'*. This is a critical point and the success of this project was supported by clearly

establishing the goals of such an investment and designing the program based on characteristics of effective professional development (Hawley and Valli, 1999).

Throughout the Australian Federal government's initiative, *Teachers for the 21st Century* (2001), reference is made to the connection between teachers' learning and student learning outcomes. However, establishing what these specific connections are, has been challenging.

Guskey's model of evaluating teachers' learning provides the opportunity to recognise that outcomes for teachers and organisations can underpin the connection to improved outcomes for students. In the project described in this paper teachers recognised their improved knowledge and skills promoted a more confident approach in their abilities to make a difference for students with learning difficulties. This would appear to be a critical connection to improved learning outcomes for students.

The project continues to provide opportunities to gather further valuable information. Future evaluation processes could include interviews with participants to investigate the continued influence of the course on their practices. Ongoing evaluation may also seek to establish whether the participants continue to engage in reflective teaching or whether they have established more formal action research projects with other colleagues. It would also be of interest to track career changes in future years. To measure outcomes for students, the participants could share their beliefs about how the course had specifically supported them to make differences for students in an ongoing way. In addition, students could also be interviewed to establish their perceptions about whether there had been changes in teaching strategies, enthusiasm and/or classroom programs. Students could be asked how these changes had been of benefit to them. For those who prefer more quantitative measures, of which there were some participants in the program, establishing pre and post program measures of student learning or using comparison groups would also provide evidence to evaluate the program at future points. However, as Guskey points out, investigators must establish where they are going if they wish to find out that they have arrived. The use of pre program desired outcomes in this project certainly allowed the participants to reflect on whether they 'got what they came for'.

In gathering planning and ongoing data, investigators need to carefully establish what they are looking for. For example, in reflecting on this project, the pre program information gathered from participants only provided information about their personal desired outcomes. Information was not specifically sought to inform the course coordinator about participants' desired outcomes in respect of influencing students' learning or school organisation, levels 3 and 5 of Guskey's model. Although making a difference for students was alluded to in the participants' pre program comments, and indeed one participant noted there were no people with training in learning difficulties at his location, seeking more specific information at all levels of Guskey's model would be a worthy consideration.

The use of Guskey's model to evaluate outcomes for participants involved in professional development programs is only one element within the bigger picture of evaluation. As noted previously, a number of people representing different organisations contributed to the design of the present program. Contributors included the South Australian Department of Education, Training and Employment (SA DETE), parent advocates, community consultants in the area of learning difficulties and university personnel. All of these contributors had a vested interest in outcomes at different levels of Guskey's model. For example, SA DETE could be interested in outcomes at all levels but in particular level 3, 4 and 5, whereas parent advocates may be more interested in evaluation at level 5 being student learning outcomes. While course aims were established based on input from these groups, this process only sought to guide the design of the program rather than clearly identify the desired outcomes of the different parties. To determine if all parties 'get what they come for' specific details would need to be gathered from these parties prior to the commencement of a program. In completing such a strategic process, evaluation beyond outcomes for participants may also be possible.

REFERENCES

Angelo, T.A. & Cross, K.P. (1993) *Classroom assessment techniques: A handbook for college teachers.* 2nd Edition. San Francisco: Jossey-Bass.

Commonwealth of Australia (DETYA) 2000 Teachers for the 21st century - Making a difference. www.detya/gov/au 6/12/2000

Elmore, R.F. & Burney, D. (1999) Investing in teachers learning: Staff development and instructional improvement. Chapter 9 in Darling-Hammond, L. & Sykes, G. (Editors) *Teaching as the learning profession: Handbook of policy and practices.* San Francisco: Jossey-Bass.

Guskey, T. R. (1999) *New persepctives on evaluating professional development.* Paper presented at the American Educational Research Association Meeting, Montreal, Canada.

Guskey, T.R. (1998) *The age of accountability.* www.nsdc.org/library/jsd/guskey194.html 6/8/2001

Guskey, T. R. (1997) Research needs to link professional development and student learning. *Journal of staff development*, 18 (2), 36-40.

Guskey, T (1995) *Research oriented professional development: In search of an optimal mix of effective practices.* www.ncrel.org/sdrs/areas/rpl_esys/pdlitrev.htm 4/6/01

Harris, A. (2000) What works in school improvement? Lessons from the field and future directions. *Educational research*, 42(1), 1-11.

Hawley, W. D. & Valli, L. (1999) The essentials of effective professional development: A new consensus. Chapter 5 in Darling-Hammond, L. & Sykes, G. (Editors) *Teaching as the learning profession: Handbook of policy and practices.* San Francisco: Jossey-Bass.

Hopkins, D (1993) A teacher's guide to classroom research. Buckingham, MK: Open University Press

Kemmis, S & McTaggart, R. (1988) (Editors) *The action research reader.* 3rd Edition. Geelong, Victoria: Deakin University Press.

Sykes, G (1999) Teacher and student learning: strengthening their connection. Chapter 6 in Darling-Hammond, L. & Sykes, G. (Editors) *Teaching as the learning profession: Handbook of policy and practices.* San Francisco: Jossey-Bass.

APPENDIX 1: GUSKEY'S MODEL OF 5 LEVELS OF PROFESSIONAL DEVELOPMENT EVALUATION (1999)

Evaluation level	What questions are addressed?	How will information be gathered?	What is measured?	How will information be used?
1. Participants' reactions	Did they like it? Was their time well spent? Did the material make sense? Will it be useful? Was the leader knowledgeable and helpful? Were the refreshments fresh and tasty? Was the room the right temperature? Were the chairs comfortable?	Questionnaires administered at the end of the session	Initial satisfaction with the experience	To improve program design and delivery
2. Participants' learning	Did the participants acquire the intended knowledge?	Pencil and paper instruments Simulations Demonstrations Portfolios	New knowledge and skills of participants	To improve program content, format and organisation
3. Organisation support and change	What was the impact on the organisation? Was implementation advocated, facilitated and supported? Were successes shared and recognised?	School records Questionnaires Structured interviews Participants' portfolios Participants' reflections	The organisation's advocacy, support, accommodation facilitation and recognition	To document and improve organisational support To inform future change efforts
4. Participants' use of new knowledge and skills	Did participants effectively apply the new knowledge & skills?	Questionnaires Structured interviews Direct observations	Degree and quality of implementation	To document and improve the implementation of the program content
5. Student learning outcomes	What was the impact on students? Did it affect student achievement? Did it influence students' physical or emotional well-being?	Students records School records Questionnaires Structured interviews Participants portfolios and reflections	Student learning outcomes: Cognitive (performance & achievement) Affective (attitudes and dispositions) Psychomotor (skills & behaviours)	To focus and improve all aspects of the program design, implementation, and follow up To demonstrate the overall impact of professional development

A predictive model identifying latent variables, which influence undergraduate student nurses' achievement in mental health nursing skills

Ian Blackman
School of Education, Flinders University *ian.blackman@flinders.edu.au*

A hypothetical model is developed to examine factors influencing nurses' achievement in a range of mental health nursing skills. The latent variables examined fall in four broad areas including the students':

- *background (gender, predominant language used, age, prior education, previous nursing experience),*
- *previous successes with other undergraduate nursing topics (physical sciences, nursing and psychosocial sciences study areas)*
- *attitudes toward mental health nursing and mental illness causation, and*
- *relationship between learning environment and achievement (consists of such factors as the type of venues used for mental health nursing experience and teacher influence in topic delivery).*

Student achievement in mental health nursing skills is estimated by measuring student's self-rated level of confidence to undertake 20 nursing related tasks (undertaken before and after a unit of mental health studies was completed by students) and scores achieved using a 50 item multiple choice tests mental health nursing. The model examining student pathways to achievement is tested using Latent Variable Partial Least Square analysis (LVPLS).

predictive model, latent variables, undergraduate student nurses achievement, mental health nursing skills

INTRODUCTION

Being able to provide health care services with qualified nurses who can maximise mental health for people in the community is one of the charters of the tertiary education sector through their undergraduate and post graduate nursing programs. This is not always an easy task based on the negative and typically stereotypical views held by the community generally about those with a mental health problem (Wolffe et al.1996), and (Brunton 1997). Additionally there have been recent structural changes to the health care system and this is happening at a time also when many western countries are closing many of their "mental hospitals" and returning long term patients with mental health problems back into the community. Arising from these communities are nursing students who are likely to have similar attitudes and outlooks towards others with a mental illness as the rest of the community. These beliefs would influence how they relate to patients with mental health problems. There are other different reasons why there are insufficient numbers of suitably qualified nurses who can meets the mental health needs of clients both in the hospital environment and within the community. It is argued that the mental health nursing profession has difficulty attracting nurses to their discipline because as nurse graduates emerge from their undergraduate programs, they are more focused on medically or surgically oriented practices (Happell and Rushworth 2000). Other factors sugges that the nurse undergraduate programs

themselves fail to deliver appropriate mental health skills development that can neither prepare graduates for their contemporary nursing roles (which includes mental health care) nor allows them to enter an appropriate pathway for advanced mental health studies (Clinton and Hazelton 2000). The number of nurses seeking to undertake post-graduate studies (and employment) in psychiatric nursing after their undergraduate work was dwindling to such an extent that the provision for effective psychiatric nursing care for patients was being undermined and the possible loss of this service area from nursing, was being advocated. This threat to the professional context of psychiatric nursing was directly linked with the lack of educational mechanisms that might attract nurses to post-graduate studies. One mechanism identified to counter this erosion was to re-assert a psychiatric nursing component at the undergraduate nursing level and to develop predictors that could be used in identifying variables which influence undergraduate nurse' achievement in mental health nursing skills.

METHOD

Participants

A sample of 183 undergraduate nursing students (commencing their third year of university studies) were targeted for the research project. Likert based questionnaires using a four point scales (strongly agree to strongly disagree and a very difficult task to a very easy task) were distributed to students. These voluntary responses were sought from students prior to them commencing a mental health module of study with related clinical practice and again were re-distributed after students had completed all clinical and theoretical requirements for the study module. This was over a thirteen week period.

Age of participating students ranged from 19-47 years of age (mean = 24 years with s.d. of 7.3 years). 162 students were female and the remaining 21 were male. 24 students used a language other than English as their first language with 159 students being native English speakers. 29 undergraduate students possessed a minor form of nursing qualification (enrolled nurse qualification), 14 students already possessing a bachelor degree other than nursing and 35 other students had partially completed some form of tertiary education. Remaining participants (115 students) had completed secondary school at year twelve level.

91% of the student cohort had not had any previous clinical contact with patients with a mental health disorder, with the remaining 17 students having a range of contact time with clients with a mental disorder (dementia usually) from1 to 8 months at the time of the survey.

More than 54% of the students had learned nothing or very little about mental health nursing at the time of the survey with only 5 % stating that had learned a great deal about mental health nursing in the past. Asked if they wanted to work in the mental health area after graduating 37% did not responds at all, 24% indicated they would not like to, 21% were undecided with the remainder indicating this career choice positively. Just under half the student group (47%) believed that learning about mental health nursing would have a positive effect on the nursing practice overall, 15% of students were undecided, with only 2 students indicating it would be deleterious to their practice. 38% of students chose not to respond to this question.

Using latent variable path analysis (LVPLS) it was hypothesised that student success in mental health nursing could be predicted by an analysis of latent and manifest variables that could influence achievement. It was further hypothesised that any variation in achievement in mental health skills would be attributed to, by these variables.

INSTRUMENT

Rating scale questionnaires were used to collect attitudinal responses from third year nursing students at two stages, before and after their unit of mental health studies. In order to meet ethical considerations for this research project, participants were advised that their names were not required in order for the survey to be completed. They were also invited to respond and

advised that the survey completion was not part of their normal academic requirements but entirely voluntary.

Forty one items or statements relating to the mental health illness and nursing interventions were generated using the learning objectives of the topic that students were expected to undertake. The focus of 11 items of the questionnaire include students attitudes toward mental health/illness and:

- management of mental illness including use of medication;
- prognosis of illnesses;
- nursing education;
- diagnostic methods;
- causation;
- impact on personal relationships;
- funding for mental health services.

The remaining 10 items of the questionnaire examined the confidence the student nurses perceived themselves to have in relation to:

- communicating effectively with patients;
- undertaking an assessment including a mental state examination of patients

  developing nursing care strategies;
- identifying appropriate patient care outcomes;
- being empathic with patients;
- teaching living skills;
- knowledge of psych-active medications;
- managing another who is aggressive;
- working within a multi-disciplinary team

With reference to figure 1, the relationships between the inner and outer variables are identified. 18 latent variables are defined by the inner circles and 38 manifest variables are represented by the outer squares.

A. Student background variables:

1. type of first language used: English or a language other than English;
2. age of the student;
3. their orientation to mental health nursing which included whether they thought a unit of study would benefit practice overall, whether they would they want to work in mental health upon graduating and if they had had an previous instruction about mental illness/health;
4. their prior education, if they already held a degree, enrolled nurse status, completed year twelve or had engaged in any prior study;
5. if any status was given toward their nursing studies or not.

B. Variables related to previous achievement within the undergraduate nursing program including their grade(s):

6. for first year nursing core topics (health assessment, foundation nursing studies and developmental health);
7. for first year psychosocial core topics (introductory sociology and communication skills);
8. for first year science topics (bio-physical science and anatomy and physiology);
9. for first year option topic (a humanities or science based option);
10. point average for all first year scores gained;
11. for second year nursing topics (law, ethics, introductory nursing theory and clinical practice);
12. for second year psychosocial core topics (health psychology and grades given by two different teacher engaged in that topic);
13. for second year science topics (pathophysiology and microbiology);
14. point average for all second year scores gained;

C. Student perception of the confidence to perform clinical skills, beliefs about mental illness causation and equity issues as they relate to people with mental illness:

15. student self rated Likert scales of confidence about skills (20 items), beliefs about illness causation and equity (prior to going out for clinical placement);
16. area of last clinical placement (hospital);
17. student self rated Likert type scales of confidence about skills (20 items), beliefs about illness causation and equity (after going out for clinical placement).

D. Achievement in mental health nursing (criteria variable)

This was estimated by compiling scores of 50% for each piece of completed work. An essay about a pre-determined mental health topic and completion of 50 item multiple choice examination featuring the following nursing problem-based scenarios:

- organic brain syndrome;
- anxiety;
- schizophrenia;
- affective disorders.

The teaching program was conducted over five weeks in which students were expected to spend four hours per week examining the following major themes:

- theoretical approaches to mental illness
- legal aspects and mental illness
- assessment techniques and nursing practice strategies for the following mental health problems:
    - organic brain syndromes
    - anxiety

- schizophrenia
- affective (mood) disorders

Clinical experience was offered to students over a 4 week period in which they could practice clinical skills accentuated during the theoretical component offered previously.

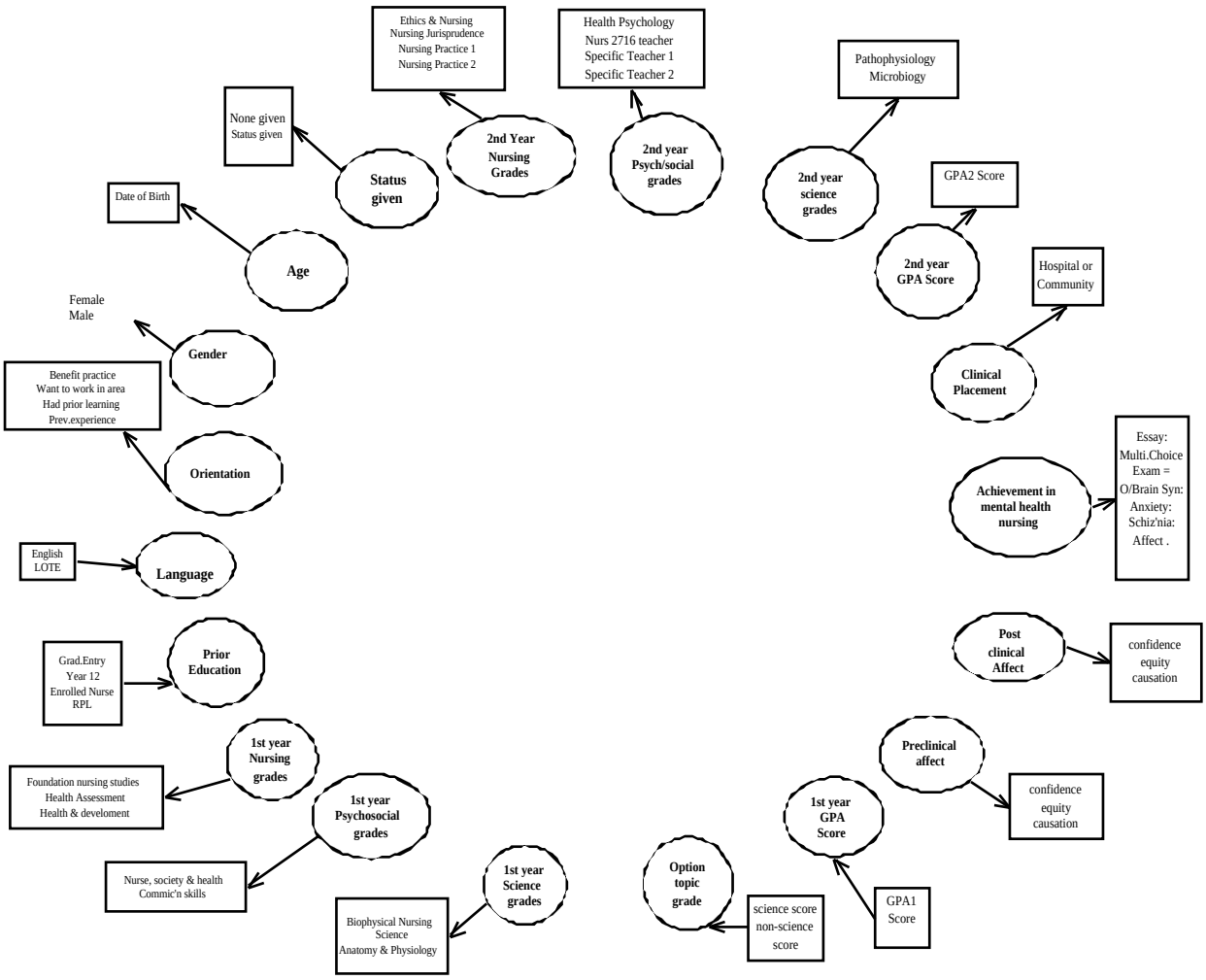


Figure 1: Hypothesised relationship between manifest and latent variables to achievement in mental health nursing

DATA ANALYSIS

Bi-variate correlation and factor analysis using SPSS was used to summarise the structure of variables related to student attitude towards patients with mental health illness. The principal components analysis in particular, was most useful for accounting for the spread or the variation of each dimension (attitude) within the multivariate space. To clarify what the student nurse attitudinal variables actually meant, varimax rotation was employed essentially to increase variable parsimony.

Then the Partial Least Squares Path (3.01) program (Sellin 1990) was used to generate the model of factors that influence achievement in mental health nursing and identified student pathways for achievement. This tool Latent Variable Partial Least Square analysis (LVPLS) is highly appropriate for analysing and predicting relationships between educational data. It is particularly helpful as it can deal with data that that are not normally distributed which would otherwise be usually excluded in traditional research approaches. The main purpose of LVPLS

is for prediction and to identify where relationships may or may not exist (Noonan, R. & Wold, H. 1985). Additionally LVPLS can be used with relative small groups of data and yet remain very robust and in situations where not all the relevant variables are known or the relationship between the latent variables (theoretical constructs) and the manifest variables are unknown or speculative (Falk 1987). LVPLS is also the modelling method of choice as it can also account for influences that traditionally confound experimental approaches. This is because it is often physically impossible to administer randomised controlled conditions to elicit causality in many educational settings (Keeves, 1988).

RESULTS

The role of language on mental health nursing success

With reference to figure 2, which examines the direct effect and indirect effects of variables of achievement in mental health nursing, it can be seen that student use of language indirectly influences achievement through the variable of orientation. Examining this relationship between language used and orientation more closely, it can be seen that a negative correlation (-1.9) exists. This suggests that those students who use English as a second language did not believe as strongly as native English speaking students, that studying a module of mental health nursing would be beneficial for their nursing practice overall.

Student age and achievement

The age of the student has a direct negative effect (-.16) on their achievement in mental health nursing. This suggests that older students are less successful that younger students. Statistical differences between the age of students was found to be significant, in that older students had greater difficulty in their ability to correctly:

- Provide nursing activity for a client who was hallucinating (p=0.008);
- Recognise symptoms/behaviours associated with thought blocking (p=0.35);
- Assess for the level of lethality associated with clients with suicide ideation (p=0.049);
- Recognise the major clinical signs of depression (p=0.001);
- Initially defuse anger in a clinical situation (p=0.006);
- Understand the rationale for provide relaxation for an anxious client (p=0.021);
- Recognised the desired effects of anxiolytic medication (0.038);
- and provide nursing activity to minimise stressors for the anxious client (p=0.009).

A negative correlation (-.15) exists between students age and orientation to mental health practice. Significantly older students were less likely to want to work in the mental health setting than younger students.

The influence of prior education and achievement

A strong negative pathway extends from age to prior education (-.52) and it suggests that it is the older students that hold more prior qualifications than the rest of the student numbers. All students who either hold a degree, have attempted but not completed some form of tertiary education or hold minor nursing qualifications are over the age of 21 years of age. Prior education also possesses an indirect effect on achievement through a through the third latent variable of orientation to mental health nursing. This suggests that both the enrolled nurse group and the students already holding a degree did not want to work in the area of mental health at the completion of their training. Prior education also holds a direct and negative pathway with achievement in mental health (-.15). This suggests that compared to the whole

student cohort, the enrolled nurse group did significantly worse with four of the five major components of the multiple-choice examination. Students holding a degree on the other hand significantly achieved better in all major components of the mental health examination than other students.

Student orientation to mental health nursing and achievement

Between the latent variable orientation and achievement, a pathway co-efficient of .29 exists suggesting that the more a student had a positive outlook and believed their studies in mental health would better prepare them for practice, the better the achievement in mental health studies. Achievement was also positively correlated to the amount of learning students had done in the past about mental health/illness.The strongest association was between student belief that learning about mental health would help their practice overall and their overall achievement in the examination (r=+.47)

Confidence about skills development, beliefs about mental illness and achievement

The pathway between latent variable called post-clinical affect and achievement in mental health has a co-efficient value of 0.1 indicating that this variable describes only 10% of the achievement variable. It should be noted however, that between post-clinical affect and the latent variable of pre-clinical affect a strong path exists; ie .45. This suggests that there is a strong link between these two variables. Closer examination shows that of the 20 items examining confidence at performing nursing tasks before clinical placement as compared to after clinical placement, more students felt more confidence in dealing with over half of the nursing practices that were asked of them. Confidence with their understanding about medication did not change much at all. Students understanding of mental illness causation changed to after clinical placement. There was significant agreement that mental illness was multifactorial in origin rather than having a single form of causation as was believed at first. Of the fourteen items that examine student beliefs about rights and equity issues associated with a person with a mental illness, there was a significant change in attitude in only four items after clinical experience. There was greater agreement that such clients should be treated equally, that they can hold responsible jobs, they should have a stronger say in how mental health funds are actually allocated and they are as trustworthy as anyone else in the community. Students were not drawn any further after clinical about whether the community is the best place where clients should be cared for or not.

Student success with second year studies and achievement in mental health nursing

Latent variable eleven which consists of an aggregate of manifest variables from student's second year of study has strong pathway to mental health nursing success (.47). In other words 47% of variance associated with achievement in mental health is explained by success in second year nursing grades. It also has strong links with three other variables arising from first year studies in science (latent variable 8), the option topic (latent variable 9) and the first year grade point average for all topic studied. It is noted that there are no direct or indirect effects of psychology (latent variable12) on the achievement in mental health nursing and the effects of placing student in clinical nursing areas (latent variable sixteen) related to mental health/illness, has negligible effects also.

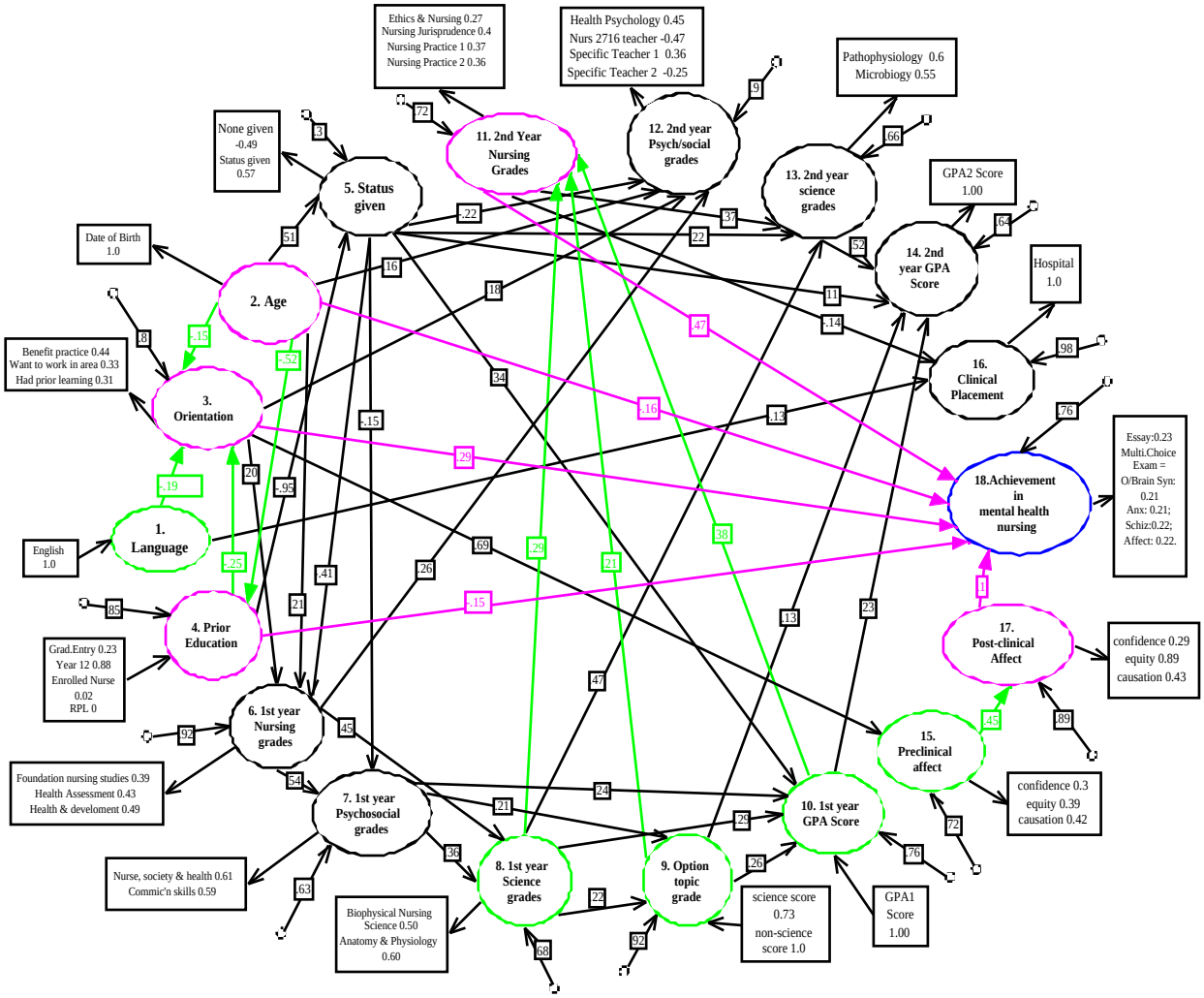


Figure 2: Final Model for Factors influencing Achievement in Mental Health Nursing

DISCUSSION

Background variables and achievement

From this study, students who use English as a second language believed there was no benefit to be derived for them by studying a topic in mental health nursing and language use was not a significant predictor for achievement in mental health nursing skills. This finding is consistent with another Australian study where confidence levels in nursing skill performance were not significantly related to language type (Bell et al. 997). In terms of achieving correct answers on the multiple choice examination, students who used English as a second language in this study showed no difference in achievement compared to native English speaking students. This finding is not supported by other studies in which it is argued that such testing can be culturally biased or too technical (Rami et al.1994) and (Byrd et al.1999). In a multicultural society like Australia, it would be important for mental health nursing preparation programs to incorporate as many nursing students who possess cultural sensitivity and who are bilingual as such nurses are potentially very positive resources to meet the various ethnic groups that present for assistance with mental health disorders.

Older students in this study, experienced more difficulty with certain parts of the student assessment requirements. These areas of concern were seen to arise around the examination process itself rather than their capacity to acquire competence in clinically related mental health

nursing skills. This finding is supported by several other studies which suggest older student performed less well academically than younger students in nurse registration courses in the US (Briscoe et al. 1999) and for nurse registration examinations (Endres 1997). With a steady increase of mature aged students accessing Australian university educational programs including nursing, mental health nursing programs need to be cognisant that many of these students bring with them different degrees of hardiness (ie: their perceived level of commitment, control and challenge) and these attributes may be quite reliable forms of predictors for success other than success with paper and pencil tests (Wright et al 19998) and (Patton et al.1999). The nursing curriculum may need to modify it processes in order to successfully incorporate the older student who is coming ibnto the nursing program to study.

One group of older students (compared to the younger secondary school undergraduates) that has demonstrated substantial achievements in mental health nursing, are those students who enter nursing already holding a degree from another discipline. These graduate entry students achieve well overall, both academically and clinically. Other studies have found similar results and it suggests that previous study at a bachelor degree level would a reliable predictor for success in mental health nursing (Barr. 1998 and Kevern et al. 1999) Greater incorporation of graduate students into mental health nursing is likely to prove to be a potentially valuable resource. It is worth noting that currently graduate entry students enter nursing at the undergraduate level and are credited with one year status from the three year nursing degree, in recognition of their already held degree. At the end of two years study graduate entry nurses exits the university with another comparable degree and need to re-enrol if they wish to go onto to take a higher degree. Alternatively, it may be better to offer a nursing program at a higher degree level (eg: at Masters level in mental health nursing) for graduate entry students in the first place as it may be more attractive to them and simultaneously meet the nursing professions' expectation to generate well-educated graduates for practice in the mental health area.

Student orientation to mental health nursing and achievement

Prior learning about mental health nursing has been seen to have a significant influence on achievement (r=+.36). Examining this relationship a little more closely, students achieved greater than 50% pass rate in sections of mental health examination paper. The most successful component of the examination paper arose from the completion of items as they relate to the nursing management of the client with organic brain syndrome i.e.: confusion. Students who have had previous experience with the aged care sector are likely to have had experience with this type of disorder. This relationship between variables highlights another important practice implication. Student nurses in this study, receive information about mental health practices in third year. Very little is given before this time. Yet nurses overall are coming into increasing contact with clients with mental disorders all the time and in health care settings not traditionally known for such interaction. With di-instituionalisation and closure of mental health facilities, clients are now seeking mental health care in community and acute care venues. At the same time, nurse's preparation for delivering specific care to these clients is delayed until late in their training. This suggests that nurses are not prepared for role in mental health nursing when in reality an expectation exists that they are able to practice in this area. Clearly achievement in mental health skills would occur more efficiently if access to mental health training was delivered much earlier within in the overall nursing curricula.

Confidence levels at completing mental health nursing skills and success in mental health examinations and tests are not good predictors for success. Of the twenty items estimating student confidence, only two had significant correlations with the overall test achievement. In terms of student beliefs about illness causation and test achievement, three were statistically significant for achievement. Student values about equity issues as they relate to clients with a mental disorder were also not good predictors for test achievement either, in that three items estimating students values were statistically significant to the examination outcomes.

Reliability of achievement in other topics as a predictor for success in mental health nursing.

Forty-seven percent of the variance in achievement in mental health nursing is explained by the variable related to student grades obtained in their second year of the training (latent variable eleven). In turn, achievement in anatomy and physiology, and biophysical sciences studied in first year explain 50% of the second year grade successes, if the path co-efficient for the first year option topic in taken into consideration. This is a indirect effect but it is substantial. Literature is divided as to the predictability that can be afforded to science for success in nursing generally however, success in nursing has been directly attributable to grades obtained in physiology and pathophysiology (Alexander et al 1997 and Byrd et al, 1999). Student's past successes in science may be unreliable for the predicting achievement in nursing outcomes, but student's self-efficacy rating for science has been cited as having a strong correlation to achievement (Caon et al 1992, Andrew et al 1998 and Ofori 2000).

The effect of the grade point average as a predictor for success in mental health nursing is not reliable as it has an indirect effect on mental health achievement. It does however have the capacity to predict achievement in the second year grades which in turn make some reasonable prediction for achievement in mental health nursing. The higher the grade point average the better the predictor for achievement in subsequent studies in nursing. (Alexander et al. 1997, Barkley et al 1998 Byrd et al 1999 and Reale, 2001). In relation to this study group, GPA scores for first year and second year was 4.9 and 4.1 respectively and these were significantly associated with mental health nursing examination scores at r =+. 27 and r =+.37 respectively.

Achievement in mental health nursing is directly associated with student success in law studies, the two nursing studies units and the achievement with the topic that examines ethics and nursing practice. All of these topics fundamentally underpin content associated with the mental health nursing studies. The relevant mental health statutes serve to inform nursing practices in mental health as does the examination of different ethical dilemmas that are associated with nursing a patient with a mental disorder. The two nursing practice topics draw on previous learning such as communication skills, sociological concepts and different nursing theory as these have a strong application to mental health nursing. The later nursing topic (nursing practice 2) takes the concepts a step further by looking at development or lifespan approaches to health and illness, as they are presented by patients both in the hospital and the community. Students are required to analyse increasingly complex clinical scenarios, which require deeper integration of pathophysiology and pharmacology. While achievement in mental health nursing skills is strongly directed by these two nursing topics however, what seems to be lacking here is the non-existent influence latent variable twelve (health psychology) is having on mental health nursing achievement. This topic serves to broaden the approach nurses take toward patient care with mental disorders as it embraces different psychological theories that apply to patients and illness. Mental health nurses need to develop a wide range of skills (and not just ones that relate to a medical model) to satisfactorily blend with the eclectic approaches that are needed in contemporary mental health nursing. From this model, the effect of psychology to actually do this is minimal, and is therefore a poor predictor for success in mental health nursing.

The effect of clinical placement (latent variable sixteen) was minimal in terms predicting achievement in mental health nursing practice. It had no direct effect on either achievement or student affect, as had been anticipated. It was believed that exposure to a mental health nursing clinical placement would assist student develop skills in promoting the nurse - client relationship, maximise their ability to assess for risks associated with mental health problems and gain first-hand experience in dealing with mental health problems. The use of diverse clinical settings particularly for experiential learning was seen to be the way to achieve this aim. While clinical placement in terms of this study was a poor predictor for success in mental

health nursing, Rhode (1996), Arnold et al. (1997) and Stevens et al. (1997) argue strongly for clinical exposure for students as a means of modifying their attitudes towards these patients.

CONCLUSION

Identifying reliable predictors for success in mental health nursing has centred on the previous achievements students have made in previous nursing or nursing related yopics. In this sample of students, prior education has a direct influence on mental health nursing achievement, particularly if the student is a graduate. Age as a predictor for success is unreliable. Use of language is not a reliable indicator either but it can influence the student's outlook about the benefits of learning about mental health. In this sample, students with a positive approach toward mental health nursing or have a desire to work in the area and who have had some prior knowledge/experience about mental health nursing are positively correlated to for achievement in mental health nursing. GPA scores inform but are not reliable indicators for success either but provide broad guidelines for estimating success. This study has also shown that nurses can be prepared for mental health nursing practice quite adequately but not necessarily with a broad perspective that in reality is often required. Their achievement in the mental health nursing topic is essentially reflective of their understanding of a dominant medical model to explain mental illness and not broadly informed by other psychological theories. Achievement in psychology studies is therefore not a reliable predictor for success in mental health nursing. From this model, success in medical and surgical nursing topics is likely to be more reliable predictor than achievement psycho-social topics.

To identify additional predictors for success in mental health nursing, further study is needed particularly taking a multi-level approach. This study has confined itself to one dimension or level by examining student attributes and their achievement to predict success in the undergraduate mental health nursing program. What is now called for is to re-examine these variables and simultaneously explore the effects of the learning institution on student achievement in mental health nursing, such as class size and the way the mental health program is delivered.

BIBLIOGRAPHY

Alexander, J. & Brophy, G. (1997) A Five-Year Study Of Graduates' Performance On NCLEX-RN. **Journal Of Nursing Education.** V36 n9 p443-445 November.

Andrew, S. & Vialle, W. (1998) Nursing Students' Self-Efficacy, Self-Regulated Learning and Academic Performance In Science. (http://www.swin.edu.au/aare/98pap/and98319.htm) University of Wollongong

Arnold, W. & Niewswiadomy, R. (1997) A Structured Communication Exercise to Reduce Nursing Students' Anxiety Prior to Entering the Psychiatric Setting. **Journal of Nursing Education**. Vol.36.N.9.pp446-447. November

Barr, J.M. (1998) Interdisciplinary Research: Predicting Academic Success In The Nursing Program. (http://www.drury.edu/faculty/ess/irconf/Jbarr.html)

Bell, A., Horsfall, J. & Goodin, B. (1997) A Multi-centre Comparative Investigation of the Impact of Undergraduate Mental Health Nursing Curricula and Clinical Experience on Attitudes, Skills and Clinical Confidence. **Report to the Department of Employment, Education, Training and Youth Affairs. Department of Behavioural and Social Sciences in Nursing, University of Sydney.**

Briscoe, V.J. & Anema, M.G. (1999) The Relationship of Academic Variables as Predictors of Success on the National Council Licensure Examination for Registered Nurses in a Selected Associate Degree Program. **ABNF Journal.** Vol.10 No4.p80-83. July-August.

Brunton, K. (1997) Stigma. **Journal of Advanced Nursing**. Vol.26.pp891-898.

Byrd, G., Garza, C. & Nieswiadomy (1999) Predictors of Successful Completion of a Baccalaureate Nursing Program. **Nurse Educator**. Vol.24. No.6. P33-37.

Clinton, M. & Hazelton, M. (2000) Scoping Mental Health Education**. Australian and New Zealand Journal of Mental Health Nursing.** V9.

Endres, D. (1997) A Comparison of Predictors of Success on NCLEX-RN for African, American, Foreign-Born and White Baccalaureate Graduates. **Journal of Nursing Education**. V36.N.8. October.

Falk (1987) **A Primer for Soft Modelling**. University of California. Berkley

Happell, B. & Rushworth, L. (2000) Can Educational Methods Influence the Popularity of Psychiatric Nursing? **Nurse Education Toda**y. V20 p318-326.

Keeves, J.P. (1988) Longitudinal Research Methods. In : **Educational Research Methodology and Measurement: An International Handbook**. Keeves, J.P. (ed) Oxford, Permagon Press

Kevern, J., Ricketts, C. & Webb, C. (1999) Pre-registration Diploma Students: A Quantitative Study Of Entry Characteristics And Course Outcomes. **Journal of Advanced Nursing.** v34 n4 p785-795

Levy, S. & Howles, K. (1995) **The Journal of Forensic Psychiatry**. Vol.6.pp19-39.

Noonan, R. & Wold, H. (1985) Partial Least Squares Path Analysis. In: **The International Encyclopaedia of Education, Research and Studies.** Husen, T. & Postlethwaite, T.N. (eds) Oxford, Permagon Press.

Ofori, R. (2000) Age and "Type" of Domain Specific Entry Qualifications as Predictors of Student Nurses' Performance in Biological, Social and Behavioural Sciences in Nursing Assessments. **Nurse Education Today.** Vol.20.ps298-310.

Patton, J. & Goldenberg, D. (1999) Hardiness and Anxiety as Predictors of Academic Success in First-Year, Full-time and Part-time RN Students. **The Journal of Continuing Education in Nursing.** V30.n4.July/August.p158-167

Rami, J. & Hansberry A. (1994) Educating Minority Students For The Health Professions: Taking A Quantum Leap To Meet The Challenge. **Education.** V115 n1 p80-87 Fall.

Rhode, D. (1996) Uncovering a World of Mental Illness: An Analysis in Perceptual Transformation in Undergraduate Nursing Students. **Archives of Psychiatric Nursing.** Vol.10.N.6.pp347-354. December

Sellin, N. (1990) **PLSPath Version 3.01 Program Manual**, Hamburg.

Setvens, J. & Dulhunty, P. (1997) A Career with Mentally ill People: An unlikely Destination for Graduates of Pre-registration Nursing Programs. **The Australian Electronic Journal of Nursing Education.** Vol.3 n.1

Wolff, G., Pathare, S., Craig, T. & Leff, J. (1996) Community Knowledge of Psychiatry and Reaction to Mentally Ill People. **British Journal of Psychiatry.** Vo.168.pp191-198.

Wright, C. , Frew, T. & Hatcher, D. (1998) Social and Demographic Characteristics of Young and Mature Aged Nursing Students in Australian Universities. **Nurse Education Today**. Vol.18.p101-107

Understanding workplace bullying: a practical application of Giddens' Structuration Theory

Rose Boucaut
University of South Australia *rose.boucaut@unisa.edu.au*

Workplace bullying is a significant and complex international issue that presents a challenge for organisations to manage. It is proposed that Giddens' Theory of Structuration (1984) may provide a basis for examining the social processes involved in the approaches adopted by organisations to manage workplace bullying. This paper presents a method for understanding organisational issues related to workplace bullying. The process of developing the individual interview questions is described. Giddens' Theory of Structuration (1984) provides the theoretical background for the questions. In 1991, Turner described Giddens' theory in terms of a framework with 11 key sensitising concepts. This paper defines each of these concepts in turn, discusses them in relation to workplace bullying and presents an interview question to explore each concept. Giddens' framework involves a series of stages, with the possibility of barriers between each of the stages. Within this theory strategies between the stages and tactics within the stages could be developed to address the problem of workplace bullying.

workplace bullying, management, Giddens' Structuration Theory

INTRODUCTION

International interest in workplace bullying has flourished over the latter part of the past decade. Concerned community groups, government agencies and workplaces have tried to discern how best to manage it; research indicates that the problem is complex and widespread. The first study in the United Kingdom to establish the incidence of workplace bullying was reported by Rayner in 1997. Her study revealed that over half of the 1137 respondents had experienced workplace bullying, and more than three quarters of respondents had observed it. These findings led Rayner to conclude that "bullying is part of many people's working lives" (Rayner, 1977, p. 207). Subsequent research in the United Kingdom by Hoel and Cooper (2000), who surveyed a representative sample of 5,288 employees at random, revealed that one in ten respondents reported they had been bullied in the last six months; nearly a quarter reported having been bullied in the last five years. Together, these studies indicate that workplace bullying is a significant hazard affecting people at work to-day. Prevention and management of workplace bullying is therefore a substantial challenge for all concerned.

Workplace bullying has various definitions; however, there are three key components to most definitions (Quine, 1999). First, it is defined in terms of its effect on the recipient not the intention of the bully; second, bullying has a negative impact on the victim; and third, it is a repeated activity. The Office of the Employee Ombudsman (1999) stated that power imbalance is a key component of all definitions, and defined bullying thus, "Workplace bullying involves the persistent ill treatment of an individual at work by one or more other persons" (Office of the Employee Ombudsman, 1999, p. 3).

Legislative provisions for dealing with workplace bullying have been called for by researchers in the United Kingdom (Rayner, 1997) and Australia (McCarthy et al, 1995; Thomson, 1997). Sweden led the world in this regard by implementing the Offensive Discrimination at Work Act 1994 (Thomson, 1997). In the United Kingdom, the Dignity at Work Bill was put before the House of Lords in 1996, aimed at preventing workplace bullying and providing "effective

remedies for the victims" (Sheehan et al, 1999). Although this Bill was not subsequently passed it was a powerful attempt to enhance harmonious workplace relationships. Currently there are moves within Australia to adopt legislative controls to address the issue of workplace bullying. In Victoria, a Draft Code of Practice was released for public comment earlier this year (Victorian WorkCover Authority, 2001) with a view to introducing a proposed Code of Practice in 2002. No formal discussion paper has yet been released in South Australia, but other public documents released locally, including research from the Working Women's Centre in South Australia (Thomson, 1997) and the Employee Ombudsman's report *Bullies Not Wanted* (1999), indicate that workplace bullying is an important issue for South Australians. This interest in workplace bullying arises from new ideas of fairness and justice that have become accepted in a changing multicultural, multiracial Australian society with equality between the sexes and across age levels as well as across different racial and cultural groups.

This paper describes the background to the method of current research aimed at understanding organisational issues related to workplace bullying in South Australia. The focus of the paper is on the development of interview questions employed in the study; these are based on Giddens' Theory of Structuration (Giddens, 1984). Research to date on organisational aspects of workplace bullying is lacking a theoretical basis, and Giddens' theory has not previously been reported in workplace bullying research literature. The following section of the paper presents information on the theory.

A DESCRIPTION OF THE KEY ELEMENTS OF STRUCTURATION THEORY

In 1984, sociologist Giddens presented his theory of Structuration (Giddens, 1984). In simple terms his theory outlines the social processes involved in the evolution of aspects of society. A key component of his theory is the double hermeneutic process, where people, upon reflection of day to day activities, are able to influence the structure of society by either reproducing current practices or by changing them. School and workplace bullying have a long history within our society and recent ideas have been advanced that challenge the appropriateness of such traditional behaviour. Turner (1991) analysed Giddens' work and produced a diagrammatic model of his Theory of Structuration. The model with its 11 sensitising concepts is illustrated in Figure 1. It is proposed in this study to use Turner's model, which consists of inter-linked but discrete concepts, to provide a framework for illuminating how a social issue, such as workplace bullying, can be managed within an organisation.

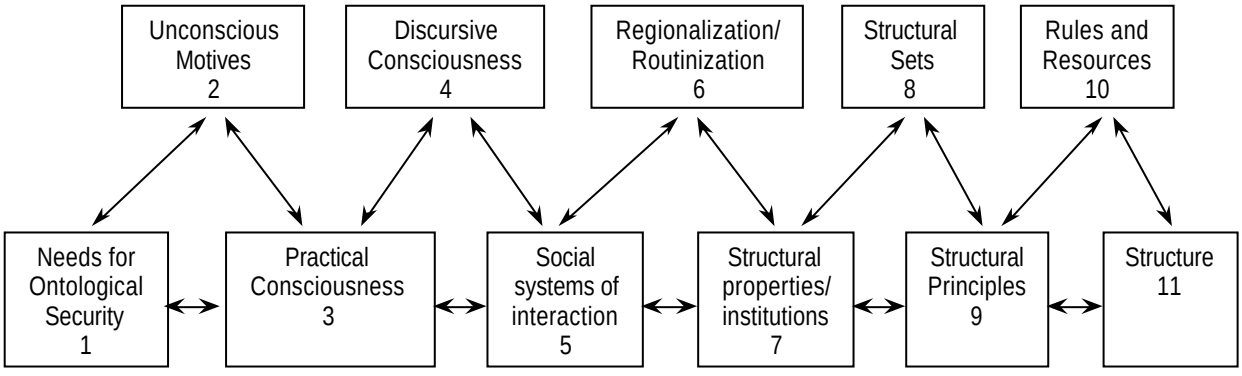


Figure 1. The 11 sensitising concepts of Giddens' Theory of Structuration presented by Turner (1991, p. 536)

Each of the sensitising concepts is defined in the following section of this paper. Where possible Giddens' glossary of terms from his work *The Constitution of Society* (1984) has been used to define the concepts. Where his definitions are unclear, Turner's reading of the concepts provides additional clarity. In each case the definition is followed by an interpretation of the sensitising concept in relation to workplace bullying. For the current study an interview question was constructed for each of these 11 concepts (shown in Table 1).

1. Needs for ontological security

Giddens described ontological security as the "Confidence or trust that the natural and social worlds are as they appear to be, including the basic existential parameters of self and social identity" (Giddens, 1984, p. 375). Turner further explained the concept,

> one of the driving but highly diffuse forces behind action is the desire to sustain ontological security or the sense of trust that comes from being able to reduce anxiety in social situations. Actors need to have this sense of trust. (Turner, 1991, p. 532)

For people to be able to perform their work to the best of their ability it is argued that they require ontological security. This is most likely to happen in a supportive work environment where the employees are able to devote their attention to the task at hand rather than being concerned for their own personal well being (mental and physical) at the workplace. The South Australian Occupational Health, Safety and Welfare Act (1986) promotes such ontological security. Under Section 19 of this Act, employers have a Duty of Care to provide a safe workplace for their staff.

2. Unconscious motives

Giddens proposed that unconscious motives play an important role in actors being able to attain ontological security. This concept was further explained by Turner, "The basic 'force' behind much action is an unconscious set of processes to gain a 'sense of trust' in interaction with others" (Turner, 1991, p. 532). These unconscious motives also lie beneath our stock knowledge of how we should act in certain situations. "There are many pressures to act in certain ways that an actor does not perceive…much of what propels action lies below consciousness" (Turner, 1991, p.532).

In the current study I aim to develop an understanding of what has led to the recent interest in the concept of workplace bullying and to tap into some of the unconscious motives that lead to the issue of workplace bullying being addressed.

3. Practical consciousness

Practical consciousness is a term that refers to "the stock of knowledge that one implicitly uses to act in situations and to interpret the actions of others. It is this knowledgeability that is constantly used, but rarely articulated, to interpret events – one's own and those of others" (Turner, 1991, p.531).

Giddens defined practical consciousness in the following way,

> What actors know (believe) about social conditions, including especially the conditions of their own action, but cannot express discursively; no bar of repression, however, protects practical consciousness as is the case with the unconscious. (Giddens, 1984, p. 375)

Each person within an organisation brings with him or her to work a certain practical consciousness or stock knowledge of what constitutes appropriate workplace behaviour. This stock knowledge forms a basis from which people undertake and interpret interactions with other staff and customers. It is possible that this stock knowledge may be modified by the requirements of the work tasks, peer pressure or the culture of the workplace.

4. Discursive consciousness

Turner states that discursive consciousness, "involves the capacity to give reasons for or rationalise what one does (and presumably to do the same for others' behaviour)" (Turner, 1991, p. 531).

Giddens defined discursive consciousness as, "What actors are able to say, or to give verbal expression to, about social conditions, including especially the conditions of their own action; awareness which has a discursive form" (Giddens, 1984, p. 374).

People within an organisation could be expected to participate in discussions about workplace behaviour. Further, people could be expected to differentiate between behaviours that are appropriate and those that are not, the latter may include workplace bullying. Similar discussions have taken place in the past within society and within organisations about other social issues, for example, harassment based on race or gender. Liefooghe and Olafsson (1999) reported that,

> One way to start increasing peoples' awareness is through discussion in groups about their understanding of what constitutes workplace bullying. This will facilitate the development of a shared frame of reference, which will help to reduce ambiguity about the interpretation of the bullying behaviour and thus facilitate preventative and remedial action. (Liefooghe and Olafsson, 1999, p. 47)

5. Social systems of interaction

Within society and within organisations there are discrete social systems of interaction. Examples of these within the general community might include religious groups and political parties. At the workplace discrete goups might include work units, for example the Research and Development Unit or the Marketing Unit, and other groups such as the staff social club. Each of these "systems" works in its own individual way.

Giddens described such systems as,

> The patterning of social relations across time-space, understood as reproduced practices. Social systems should be regarded as widely variable in terms of the degree of 'systemness' they display and rarely have the sort of internal unity which may be found in physical and biological systems. (Giddens, 1984, p. 377*)*

Given the variability that might be displayed within these units or systems it is likely that some social systems within society are more or less tolerant of workplace bullying than other systems. Similarly individual units and groups within organisations might vary in their tolerance of workplace bullying.

6. Regionalisation/routinisation

Routines that people engage in and the territory they occupy are important in establishing and maintaining relationships at the workplace. Turner writes:

> Both the ontological security of agents and the institutionalization of structures in time and space depend upon routinized and regionalized interaction among actors. Routinization of interaction patterns is what gives them continuity across time, thereby reproducing structure (rules and resources) and structures (institutions). At the same time, routinization gives predictability to actions and, in so doing, provides for a sense of ontological security. Thus routines become critical for the most basic aspects of structure and human agency. Similarly, regionalization orders action in space by positioning actors in places vis-à-vis one another and by circumscribing how they are to present themselves and act. As with routines, the regionalization of interaction is essential to the sustenance of broader structural patterns and ontological security of actors, because it orders people's interactions in space and time, which in turn reproduces structures and meets an agent's need for ontological security. (Turner, 1991, p. 533)

When people work within an organisation they become socialised into the routines of that workplace. An interruption of these routines, for example when introducing change, may cause discomfort to a person's ontological security. McCarthy et al (1995) found that during organisational restructuring, managers adopted inappropriately coercive behaviours and many employees reported being bullied.

In addition to the routines people become familiar with at the workplace, people are usually given their own workspace within an organisation. The workspace may be a place on the production line, a company vehicle, or an office. There are also generally designated areas for staff to take their breaks such as a staff tea-room. These spaces and facilities for staff fall under the umbrella of regionalisation as described by Giddens and Turner. Such distanciation enables people to maintain their own "space" and may to some extent enhance their

ontological security. Encroaching on a person's physical or mental "space" at the workplace, and causing them distress when doing this, may constitute workplace bullying.

7. Structural properties/institutions

Giddens described structural properties as "Structured features of social systems, especially institutionalized features, stretching across time and space" (Giddens, 1984, p. 377).

I will illustrate my interpretation of the term structural properties by example. In many organisations staff are consulted about the generation of a policy (discursive consciousness) and are given the opportunity to provide feedback while the policy is in draft form. Such feedback is then generally incorporated, or at least considered, before the policy is released in its final form. At the time of the introduction of a new policy, a review date is generally set at which time the policy is again circulated for comment and modification as appropriate. In this example, structure is a feature of the process; the system of consultation is sequential and crosses all sectors of the organisation.

8. Structural sets

Turner described structural sets as the "rule/resource bundles, or combinations and configurations of rules and resources, which are used to produce and reproduce certain types and forms of social relations across time and space" (Turner, 1991, p. 529). He continued, "…the general idea [is] that the general structural principles of class societies are transformed into more specific sets of rules and resources that agents use to mediate social relations" (Turner, 1991, p. 529).

For the purposes of the current investigation I have interpreted structural sets as rules and resources emanating from an organisation or from within individual units of an organisation, such as the Human Resources Unit or the Occupational Health and Safety Unit of a large organisation. Each of these individual Units would generate rules and policies that guide social relations within the organisation. For example the Human Resources Unit may develop a code of conduct or a no bullying policy and accompanying grievance procedures.

9. Structural principles

Turner has written structural principles "are the most general principles that guide the organization of societal totalities" (Turner, 1991, p. 528). Giddens defined them as, "Principles of organization of societal totalities; factors involved in the overall institutional alignment of a society or type of society" (Giddens, 1984, p. 376).

I have interpreted structural principles in several ways. I think structural principles are, in part, organisational objectives, and that they reflect the purpose of the business, such as customer service and/or a quality product. Coupled with this interpretation of structural principles as organisational objectives, is the suggestion that structural principles also include the values that an organisation holds. In this way I think that structural principles are tied to organisational culture. Schein, in his text on organisational culture and leadership stated, "that culture is a *deep* phenomenon, that culture is *complex* and difficult to understand, but that the effort to understand it is worthwhile because much of the mysterious and irrational in organisations suddenly becomes clear when we do understand it" (Schein, 1985, p. 5). Giddens suggested that contradictions in structural principles may occur. Contradictions may occur at a workplace where there is some variance between the organisational objectives and the organisational culture, and such variance may lead to ontological insecurity in employees.

A structural principle in relation to workplace bullying may be that an organisation aims to become a bully-free workplace of its own volition. The principle in this case is to deal with the issue in a pro-active manner as some companies have done (Crawford, 1999). Another workplace may decide that they would not do anything about workplace bullying unless they

were required to, for example to comply with legislation; thus their structural principle is reactive rather than proactive.

10. Rules and resources

Turner has argued that "Rules are 'generalizable procedures' that actors understand and use in various circumstances" (Turner, 1991, p. 523). Organisations themselves abide by the rules of the society in which they are located and each individual organisation has its own rules. In South Australia employers are required to comply with legislative requirements of the Occupational Health, Safety and Welfare Act (1986). Turner goes on to say "Resources are facilities that actors use to get things done" (Turner 1991, p. 524), and "Rules and resources are mediating in that they are what tie social relations together. They are what actors use to create, sustain, or transform relations across time and in space. And, because rules and resources are inherently transformational – that is, generative of diverse combinations – they can lace together many different patterns of social relations in time and space" (Turner 1991, pp. 534-535).

Within all organisations there are explicit and implicit rules that guide peoples' behaviour. In large organisations explicit rules might include a policy relating to behaviour for example regarding bullying at work. Implicit rules is a term used to describe the generally unspoken and unwritten rules of the workplace (Hopkins 2000). These rules might include some of the initiation rituals that workers sometimes use to "welcome" new staff.

Resources that might be mobilized in putting a no-bullying policy in place include staff time for the development of policies and procedures, and resources allocated for educating staff, including down time to attend training sessions.

11. Structure

Giddens defined structure in the following manner, "Rules and resources, recursively implicated in the reproduction of social systems. Structure exists only as memory traces, the organic basis of human knowledgeability, and instantiated in action" (Giddens, 1984, p. 377).

For the purposes of the current study I have interpreted structure to mean the management structure of an organisation, such as a flat or hierarchical management system. An organisation with a hierarchical management structure, such as the fire service, may provide particular opportunities for workplace bullying (Archer, 1999).

The double hermeneutic principle

Structural change in organisations is likely to occur in response to pressures exerted by forces external to, and/or from within an organisation. In workplace bullying such external pressures, may, for example, be new ideas circulating in the wider community that lead to legislative requirements which arise in response to lobbying from community groups, unions or regulatory agencies. Internal forces trying to address workplace bullying may arise at various levels within the organisational structure, for example, from senior management, from units concerned with staff welfare such as occupational health and safety or human resources, or from groups of concerned workers. Giddens describes a process called the double hermeneutic where people can bring about change through their actions following reflection on their own world view and in the light of new knowledge and understanding. The internet is a particularly useful vehicle for both researchers and lay people to access new material and the rapid speed of information transmission enables people interested in the subject of workplace bullying to keep up to date and informed about current issues and research. Equality of opportunity for women in the workplace was the focus of public discussion in the past and following this people across the globe are now striving for equity and fairness at the workplace. The double hermeneutic principle underlies this process of new knowledge influencing social and structural change.

THE AIMS OF THE CURRENT STUDY

The current study seeks to answer the three following research questions:

1. Why has workplace bullying become an important issue to South Australian organisations?
2. What are the stages and barriers to addressing the issue of workplace bullying?
3. What strategies and tactics are employed in addressing the problem of workplace bullying?

The method used to answer these questions will be to interview a purposeful sample of key stakeholders within the movement against workplace bullying in South Australia. The purpose of such detailed consideration of Giddens Theory of Structuration and Turner's reading of the Giddens work is to investigate whether Turner's framework is useful in exploring the concept of how organisations manage workplace bullying. Defining and interpreting the 11 sensitising concepts described by Turner (1991) is undertaken with the deliberate purpose of composing interview questions to examine the concepts. In order to do this a series of interview questions (Table 1) have been developed to illuminate each of the sensitising concepts described by Turner (1991). These questions are listed in Table 1 and are aligned with Turner's 11 sensitising concepts that are illustrated as cells in Figure 1. The people in the purposeful sample hold positions where they provide "expert advice" to organisations about the management of workplace bullying. They therefore have an important educative role in assisting organisations to learn about, and deal with, the problem of workplace bullying. Three pilot interviews have been conducted prior to the study. The focus of the first two was on the research questions, the focus of the third was on the interview questions; these seemed to work reasonably well, although a slight change in the sequence of the questions was decided upon for the actual study. To date, four interviews with experts have been conducted, another three are scheduled. In each case respondents have been sent the questions at least a week in advance of the interview.

*Table 1. Interview questions aligned with Turner's (1991) sensitising concepts based on Giddens' Theory of Structuration (Giddens, 1984)**

1	What conditions must exist within organisations for people to feel free of bullying at the workplace?
2	How did the issue of workplace bullying come to the attention of South Australian organisations?
3	What sort of issues regarding workplace bullying are organisations concerned about?
4	How do staff within organisations go about raising issues related to bullying that they want aired or addressed?
5	What are the particular ways of behaving towards each other at workplaces that have been set up that encourage/discourage bullying?
6	What routine procedures/strategies do organisations put into place to try and promote a bully-free workplace?
	What spaces have organisations set up to promote a bully-free workplace?
7	What organisational systems enable (or provide barriers to) the bullying issue being addressed?
8	How do individual organisations work in different ways to promote a bully-free environment?
9	What are the organisational principles and values that encourage/discourage bullying at work?
10	What explicit rules are there (if any) regarding bullying behaviour within South Australian organisations?
	What implicit rules are there (if any) regarding bullying behaviour within South Australian organisations?
	What resources have organisations committed to addressing workplace bullying?
11	How do you see that an organisations' management structure may influence bullying at the workplace?

* Question numbers in the left hand column correspond with cell numbers in Figure 1

REFLECTIONS AND DISCUSSION

How to address the problem of workplace bullying is an issue of international interest due to the extent of the problem (Hoel and Cooper, unpublished; McCarthy et al, 1995; Rayner, 1997;

Thomson, 1997). The adverse effects on the victim may be physical, emotional or behavioural (Lewis, 1999; Quine, 1999; Thomson, 1997). There is considerable expense to organisations arising from workplace bullying. Direct costs include medical and legal costs (McCarthy, 1999), indirect costs include dysfunction and inefficiency at the workplace. For example, Rayner (1997) found staff turnover due to workplace bullying may be as high as 25%. Ultimately the community pays a high price for workplace bullying.

This paper has described the background to the interview questions I have developed for my current research into organisational aspects of workplace bullying. To my knowledge Giddens' Theory of Structuration has not previously been used as a basis for workplace bullying research. Turner's (1991) model should prove useful to practitioners working with organisations to address workplace bullying. It illustrates the sequential stages that an organisation might consider when working to manage the problem. These stages start with recognising that an individual feels unsafe when bullied at the workplace (ontological security), and go to discussions about workplace bullying within the organisation (discursive consciousness), to decisions about whether to tackle bullying in a proactive or reactive manner (structural principles). Further stages include the implementation of policies (rules and resources) and an examination of the management set-up itself (structure). The model may also be helpful in anticipating where barriers to managing workplace bullying will arise so that thought can be given as how best to overcome them.

A challenge I faced in using Giddens' theory was defining each of the 11 sensitising concepts described by Turner (1991). Some of the definitions provided by Giddens are unclear and even after having sought further understanding through Turner's reading of Giddens model, they remain unclear. In those cases I interpreted them for the purposes of the study. One such difficulty arose in my own interpretation of the difference between structural properties and structural principles and I did not find my reading of either Giddens' or Turners' work really helped differentiate and clarify the differences for me. In order to overcome this sort of difficulty I spent considerable time in round table discussions with two senior academics, one familiar with Giddens' and Turners' work and the other familiar with the organisational processes involved in addressing bullying. Together we clarified our understanding of the theory.

It is not known at this stage how clearly respondents' responses will illustrate the sensitising concepts. From the pilot interviews conducted to date, it would seem that some of the concepts such as ontological security, discursive consciousness, and rules and resources are all themes that respondents can clearly describe. One concept that some pilot respondents found difficult to relate to is the concept of regionalisation (spaces) and routinisation, particularly the implicit rules.

There is a considerable amount of overlap between the 11 sensitising concepts and this is illustrated by the double headed arrows in Figure 1. This meant that in some of the interviews conducted to date there has been some repetition in responses as the interview progressed. Whether this would have occurred if the interview questions had been asked in the reverse order, that is starting with the questions about structure and working through to ontological security is not known. As victim vulnerability is, by definition, a feature of workplace bullying it seemed logical to start the interview off with a question about ontological security.

The responses respondents give will provide information to answer the three research questions of why the issue has become important, barriers to addressing the problem and factors that constitute best practice in managing workplace bullying at an organisational level.

CONCLUSION

Workplace bullying is a serious international problem that is bereft of a theoretical basis for understanding it as an organisational problem. Research and interview questions have been presented that are currently being used in a South Australian study to further an understanding

of the organisational aspects of managing workplace bullying. The interview questions have been developed based on Giddens' Theory of Structuration (1984). From the interpretation of Giddens' theory provided by Turner (1991), a model with 11 sensitising concepts has proved to be useful in developing a grasp of the theory. These concepts were considered in the light of workplace bullying within an organisation and a question was developed for each concept. The development of questions following the model gives a structured basis to the interviews. Responses should provide a deeper understanding of how organisations manage workplace bullying; information that is likely to be useful to practitioners working in this field. Challenges and changes to existing and accepted workplace practices reflect the double hermeneutic process at work.

REFERENCES

Archer, D. (1999) Exploring "bullying" culture in the para-military organisation. *International Journal of Manpower*, 20 (1/2), 94-105.

Crawford, N (1999) Conundrums and confusion in organisaions: the etymology of the word "bully". *International Journal of Manpower*, 20 (1/2), 86-93.

Giddens, A. (1984) *The Constitution of Society.* Cambridge*:* Polity Press.

Hoel, H. and Cooper, C. *Destructive Conflict and Bullying at Work*. Unpublished report sponsored by the British Occupational Health Research Foundation. Manchester School of Management, University of Manchester Institute of Science and Technology, 2000.

Hopkins, A. (2000) *Lessons from Longford: The Esso Gas Plant Explosion.* Sydney: CCH Australia Limited.

Lewis, D. (1999) Workplace bullying – interim findings of a study in further and higher education in Wales. *International Journal of Manpower*, 20 (1/2), 106-118.

Liefooghe, A. and Olafsson, R. (1999) "Scientists" and "amateurs": mapping the bullying domain. *International Journal of Manpower*, 20(1/2), 39 – 49.

McCarthy, P., Sheehan, M. and Kearns, D. (1995) *Managerial Styles and their Effects on Employees' Health and Well-being in Organisations Undergoing Restructuring*. Brisbane: Griffith University.

McCarthy, P. (1999) Bullying at work: implementing industry self-regulation guidelines. *Journal of Occupational Health and Safety - Australia and New Zealand*, 15(5), 417-422.

Occupational Health, Safety and Welfare Act (1986).

Office of the Employee Ombudsman, (1999) *Bullies Not Wanted*. Adelaide: Office of the Employee Ombudsman.

Quine, L. (1999) Workplace bullying in NHS community trust: staff questionnaire survey. *British Medical Journal*, 318 (7187), 228-232.

Rayner, C. (1997) The incidence of workplace bullying. *Journal of Community and Applied Social Psychology*, 7(7), 199-208.

Schein, E.H. (1985) *Organizational Culture and Leadership*. San Francisco: Jossey-Bass.

Sheehan, M., Barker, M. and Rayner, C. (1999) Applying strategies for dealing with workplace bullying. *International Journal of Manpower*, 20 (1/2), 50-56.

Thomson, C. (1997) *Workplace Bullying Project*. Adelaide: Working Women's Centre of South Australia Inc.

Turner, J.H. (1991) Structuration Theory of Anthony Giddens. In *The Structure of Sociological Theory* (5th ed). California: Wadsworth, pp. 519 – 539.

Victorian WorkCover Authority, (2001) *Issues Paper: Code of Practice for Prevention of Workplace Bullying*. Melbourne: Victorian WorkCover Authority.

Acknowledgements

I would like to thank Professor John Keeves and Associate Professor Phillip Slee from Flinders University for their assistance.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Assembling an item-bank for computerised linear and adaptive testing in Geography

Kristian Leonard Burghof
kburghof@aol.com

An item-bank is a relatively large collection of easily accessible items stored in a database, categorised and coded by topic, year level, item type, graphic file source, thinking skills and processes from Blooms Taxonomy and item difficulty. This paper summarises the procedures used in planning, assembling, calibrating and entering items into a database for an item-bank for computerised linear and adaptive testing in Geography.

The process of item banking took approximately six months to assemble 152 multiple-choice items (questions) relating to Geography used to assess a Year 8/9 topic on maps and mapping skills. It is also the first time a calibrated item-bank has been assembled for computerised linear and adaptive testing in Geography, since other areas such as Mathematics, Science (e.g. Physics and Chemistry), English and Languages Other Than English (LOTE) (e.g. Spanish, French and German) are currently available. Some exemplars of items are highlighted.

item-bank, computerised linear on-line testing (LOT), computerised adaptive testing (CAT), item calibration, database

INTRODUCTION

Teachers spend incessant hours planning for tests, assembling items, writing and evaluating them, determining the difficulty level of each item (Rudner, 1998b) and then making photocopies of the test on paper for each student. Why not have a bank that stores test items, that can be reused, are easily accessible, durable and ready at an instant for computerised testing. In the study by Burghof (2001), titled "Computerised Adaptive Testing (CAT) in Society and Environment (SOSE): An exploratory study in Year 8/9 Geography, maps and mapping skills" an item-bank was assembled for computerised linear on-line and adaptive testing in Geography. Without an item-bank for the study, it would not be possible for a student to sit for a computerised test. An item-bank allows for a test to be long or short, easy or hard, depending upon the purpose of test. During a test, items are drawn from a bank one at a time (e.g. adaptive tests) or selected by the teacher prior to the test (e.g. linear tests), which are then administered to students on computer. At the end of a computerised test, immediate accurate scoring and feedback is given to all students (see Burghof, 2001).

ASSEMBLING AN ITEM-BANK FOR COMPUTERISED TESTING

There are two forms of computerised tests that require an item-bank, the first is linear tests and the second is adaptive tests. As the name suggests, both tests are conducted on computer and can be administered on-line (Internet). Computerised Linear On-line tests or LOT for short are computerised forms of the traditional pencil and paper test (see Burghof, 2001). During a test, students are presented with a set number of items (e.g. 30 items) that cover a broad range of difficulty levels (e.g. east to hard). Questions are presented to students in a linear format, one question after another going down the page. The administrator or teacher selects the items prior to testing and in most cases, little consideration is given on the part of the teacher about a student's performance on previous questions (TOEFL, 2001). On the other hand, a

computerised adaptive test or CAT for short tailors items to student's ability level. During the test, the computer will present one item to the student, accept and score a students response given, re-estimates the student's ability level and selects a new item to be administered from an item-bank (Burghof, 2001). CAT will only stop when a student's ability level has been estimated.

WHAT ARE THE BENEFITS OF ITEM BANKING?

Some advantages to item banking include flexibility, security and consistency (Umar, 1999). Rudner (1998b) has indicated that items in an item-bank can be edited and withdrawn and populated with new items when needed. Rudner (1998b) further suggests that items can be used to diagnose tasks that students can and can not do, since items located on the Rasch scale allows a teacher to identify the difficulty of items given to students.

WHAT ARE THE CHALLENGES TO ITEM BANKING?

There are some challenges to developing an item bank for computerised linear on-line and adaptive testing. Firstly, it takes much time to plan and assemble items for an item-bank, to calibrate and then enter them into a database. This process required patience and much computer expertise (Rudner, 1998b). Finally, items need to be continuously reviewed and calibrated.

PLANNING AND ASSEMBLING ITEMS

There are several crucial steps that need to be considered when planning and assembling items for an item-bank for computerised linear on-line and adaptive testing, these are as follows (see Burghof, 2001).

Drawing a Blue Print – Teachers' Experiences

The initial step in item banking is to draw a blue print of the topic from teachers' experiences. In order for this to be carried out, the subtopics need to be identified and clarified. For example, let us assume that four Geography teachers were asked to draw a blue print of a unit of work on maps and mapping skills for a Year 8/9 class. Each teacher would be asked to clearly indicate the subtopics they would teach for the topic and the learning objectives for each subtopic. Care needs to be taken when collating the subtopics from the teachers' blue prints into a unified unit of work on maps and mapping skills.

Table 1 indicates a hypothetical school, were four teachers have identified the subtopics they would include in a unit of work on the topic. The first column indicates the subtopics collated from the teachers' blue prints into 13 subtopics. The next column indicates the four teachers' selection of each subtopic (*T1, T2, T3 and T4*), and the final column indicates the number of responses given for each subtopic (*Ntm*). In this case, 13 subtopics have been indicated as possible areas for a unit of work on the topic. Factors such as curriculum and school resources (e.g. library, computer facilities) and time may influence a teachers choice of subtopics to be taught in the classroom.

Gathering Resources

Once the subtopics have been identified, resources need to be gathered and learning objectives then need to be clarified. Gathering resources on each subtopic may help clarify the learning objectives for each subtopic. Types of resources sought may include class textbooks, videos, teacher references, Geography Internet Web pages and computer software applications. Once the subtopics and learning objectives have been identified, they need to be validated (content validity) by known people in the field the item-bank is to be developed, such as Geography teachers.

Table 1. Blue print of teachers' perspective's of a unit of work on maps and mapping skills

Subtopics	*T1*	*T2*	*T3*	*T4*	*Ntm*
1. Symbols on Maps	✓	✓	✓	✓	4
2. Directions on Maps	✓	✓	✓	✓	4
3. Reading Contour Lines on a Map		✓	✓	✓	3
4. Contour Lines and Patterns		✓	✓	✓	3
5. Language of Maps	✓			✓	2
6. Latitude and Longitude	✓	✓	✓	✓	4
7. Grid and Area References	✓	✓	✓	✓	4
8. Grid Lines	✓	✓	✓	✓	4
9. Atlas Index Skills			✓	✓	2
10. Abbreviations in an Atlas			✓	✓	2
11. Time Zones	✓	✓	✓	✓	4
12. Scales on Maps	✓	✓	✓	✓	4
13. The World's Continents and Oceans		✓	✓	✓	3
Total	8	10	12	13	43

The next step is to select, scan, save and edit appropriate graphics from the resources previously gathered. For example, various computer technologies, such as a scanner can be used to scan graphics form pages in textbooks, and save them in graphic format like jpeg. However, difficulties may arise when scanning graphics from A3 textbooks (e.g. doubled paged graphics), because a scanner can only capture one A4 page at a given time. To overcome this problem, figures can be scanned in several parts and then modified in a graphics application (e.g. *Microsoft Paint*), where the graphic can be manipulated back into its whole graphic. A second difficulty may arise when graphics that are saved in jpeg or gif format are too large for use in testing. *Microsoft Image Composer Version 1.5* can be used to reduce the size of the graphic or table to suitable sizes for testing. In some cases, a graphic that has been reduced in size may result in the text to become unreadable (e.g. font size becomes too small). Again, a graphics application can be used to edit the graphic file, re-type any text in a larger font size so a test taker can easily read it. Finally, to avoid copyright and plagiarism, due reference should be given to all resources used to create items for the bank. Each item should have a reference for the question, graphic and/or table and multiple-choice options.

Estimating the Number of Items for the Bank

The next stage is to make an estimation of the number of items to be included in the bank. Wood and Skurnik (1969) recommend a total of 200 or more items for an item-bank. However, Sands et al. (1997) indicates that an item-bank should contain a ratio of items to the number of test takers. For example, 5 to 10: 1, where 5=minimum number of items, 10=maximum number of items and 1=number of test takers. In other words, 5 to 10 items are required for 1 test taker. One of the purposes for this is to avoid item repetition during testing. For the purpose of this paper, the ratio of items to the number of test taker's method is used (see Sands et al. 1997). This method was chosen for its flexibility when designing items within a particular range group (e.g. minimum to maximum) instead of one designated number such as 200. In order to calculate the ratio of items to the number of test taker's, it is necessary to estimate or have knowledge of the number of students to be tested. In this case, let us assume

there are 25 students (n=25) who will take a test on computer. To accommodate for the number of subtopics selected in Table 1 for the item-bank, a mathematical formula was created and designated as item bank ratio. To identify the ratio of items to the number of test taker's, the following two-part (A and B) mathematical formula was used: **Item Bank Ratio = Min to Max: Ns**.

PART A PART B

$$\textbf{Item Bank Ratio} = \textbf{Min (Min x Ns) to Max (Max x Ns): Ns} \qquad \text{(Eq.1)}$$

Where, Min=minimum number of items required for an item bank, Max=maximum number of items required for an item bank and Ns=number of students in class.

- To calculate the Min value (Part A). To calculate the item bank ratio of items, it is necessary to identify the minimum number of items required for the bank (Min). Let us assume the Min=5 and Ns=25. To calculate the Min value, we need to multiple the Min=5 by the Ns=25. This will give a Min value of 125, which means 125 is the minimum number of items required for the bank.
- To calculate the Max value (Part B). Similarly, to calculate the Max value, it is necessary to identify the maximum number of items required for the bank. Let us assume the Max=10. The number of test taker's (Ns) stays the same as above. To calculate the Max value, we need to multiple Max=10 by Ns=25. This will give a Max value of 250, which means the maximum number of items for the bank is 250. Thus, the formula: Item Bank Ratio = Min (125) to Max (250): Ns (25). In other words, 125 to 250 items are needed for 25 test takers.

When assembling an item-bank for computerised linear on-line and adaptive testing it is important to have an estimate of the number of items required for the bank. However, it is not necessary to estimate the right distribution of items for each subtopic in computerised adaptive testing. CAT required items to be calibrated to their difficulty level along a common scale (Rasch scale). This will be discusses later in the paper. The next section only applies for computerised linear on-line tests (LOT).

Estimating the Right Distribution of Items for each Subtopic

For LOT it is required to estimate the right distribution of items for each subtopic. A mathematical formula by Wood and Skurnik (1969) was adopted and modified for use in this paper to identify the individual subtopic requisite ratio (designated as ItRn) for each subtopic. The formula used: **ItRn = MinA to MinB: Ns**.

PART A: MinA PART B: MaxB

$$\textbf{ItRn} = \frac{\textbf{Ntm}}{\textbf{M}} \times \textbf{Min} \quad \text{to} \quad \frac{\textbf{Ntm}}{\textbf{M}} \times \textbf{Max}: \textbf{Ns} \qquad \text{(Eq.2)}$$

Where, Ntm=number of times the topic was mentioned, M=total number of mentioned of all topics, Min=minimum number of items required for an item bank, Max=maximum number of items required for an item bank and Ns=number of students in class. To identify the ItRn it is necessary to use the data in Table 1 and the Min and Max values calculated previously in this paper. As an example, let us use the subtopic Symbols on Maps from Table 1.

- To calculate the minimum (MinA) number of items to assemble for each subtopic, it is necessary to take the number of times the subtopic was mentioned by the teachers (Ntm=4) and multiple this by the Min value (Min=125) calculated previously. Then, we need to divide this by the number of responses mentioned of all topics (M=43). The formula for PartA may look like this: ItRn=(4x125/43). This will give a value for MinA=12, which means a minimum number of items for the subtopic Symbols on Maps is 12.
- Similarly, to calculate the maximum (MaxB) number of items to assemble for Symbols on Maps it is necessary to take Ntm=4 and multiple this by Max=250, which was calculated previously. Then, we need to divide this by M=43. The formula for partB may look like this: ItRn=(4x250/43). This will give a value for MaxB=23, which means a maximum number of items for the subtopic Symbols on Maps is 23. In this case, the ItRn=12 to 23: 25. In other words, 12 to 23 items need to be created for the subtopic for 25 students. Table 2 indicates the ItRn for each subtopic used as an example in this paper. In the first column it shows the 13 subtopics identified by the four teachers. In the second column are the MinA values and the third column indicates the MinB values. For example, the subtopic Language requires 6 to 12items to be created for 25 students in the class (ItRn=6 to 12: 25).

Writing Multiple-Choice Items

With the required number of items for an item-bank established it is necessary to begin to design and write multiple-choice items for the bank. A multiple-choice question or MCQ for short can be a question or incomplete statement, which asks a test taker to select one response from a list of possible alternatives (Frary, 1995; McMillan, 1997; Marsh, 1998; Theobald, 1974). The question is known as the stem, the answer is known as the key and any undesired responses are called distractors. There are several rules that need to be followed when writing multiple-choice items these are as follows.

1. Items need to assess a particular aspect of the topic.
2. Items should be independent of one another. To answer the question, all information should be found in the question.
3. The question should be written in such a way to avoid ambiguity.
4. Distractors should be plausible. They should attract test takers with little or no knowledge of the question away from the answer or key.
5. Response options should be written to equal length, as not to make it possible for a test taker to identify a correct or incorrect option from its length.
6. When a multiple-choice item is written as an incomplete statement, distractors should follow gramatically from the question.
7. The question shouldn't include negative words such as except and not. This may confuse, frustrate or cause anxiety for test takers.
8. Use "none of the above" as a final response option (e.g. for computations).

As an aid in writing multiple-choice items, Blooms Taxonomy can be used to categorise items into different cognitive domains that assess different thinking skills and processes of students (Granello, 2001; McMillan, 1997; Usova, 1997). There are six thinking skills and processes, these include knowledge (remember and recall), comprehension (understand), application (execute and implement), analysis (differentiate and organise), synthesis (plan, produce and create), and evaluation (check and critique).

Table 2. Estimating the number of items for each Subtopic

Subtopics	MinA	MinB
1. Symbols on Maps	12	23
2. Directions on Maps	12	23
3. Reading Contour Lines on a Map	9	17
4. Contour Lines and Patterns	9	17
5. Language of Maps	6	12
6. Latitude and Longitude	12	23
7. Grid and Area References	12	23
8. Grid Lines	12	23
9. Atlas Index Skills	6	12
10. Abbreviations in an Atlas	6	12
11. Time Zones	12	23
12. Scales on Maps	12	23
13. The World's Continents and Oceans	9	17
Total	125	250

When writing items for an item-bank, Theobald (1974) and Wood and Skurnik (1969) recommended the use of item cards for ease of classification and editing. An item card is a single A4 page, and written on it is a single multiple-choice question, stimulus material (graphic or table), response options, references for the item and item bank number.

Once items have been written on item cards, it is necessary to have them validated for subject content and for Blooms Taxonomy. Items should be checked for year level suitability, repetition and if items have been categorised into the correct cognitive domain from Blooms Taxonomy.

CALIBRATING ITEMS FOR CAT & LOT

This section will briefly describe the Rasch Model, common-item equating procedures and item calibration using *ConQuest*.

The Rasch Model

When developing an item-bank, items can be calibrated to a common measurement scale (Rasch scale) also called a log-linear scale. The Rasch model is based on the probability that a person with a given ability will answer correctly an item representing a given difficulty, as shown in Figure 1 (Wright and Stone, 1979). Figure 1 shows the relationship between item difficulty and person ability. The unit of measure on the log-linear scale is as a logit (Keeves and Alagumalai, 1999; Wright and Stone, 1979), and items will generally fall between –4 and +4 logits (the mean of the scale is 0). Items with a negative logit value (e.g. -3) tend to be easy, while an item with a positive logit value (e.g. +3) tends to be difficult.

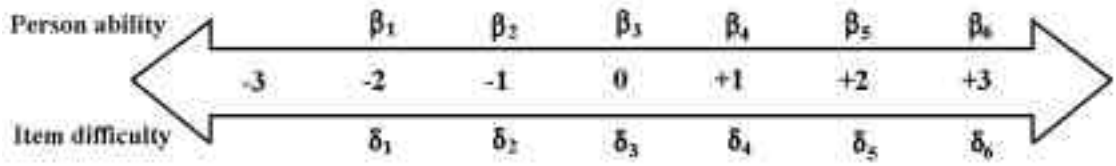


Figure 1. The Relationship of Item Difficulty to Person Ability (After Keeves and Alagumalai, 1999, Figure 1, p. 27)

Common Item-Equating

In common-item equating, two tests (A & B) are administered to two separate groups (A & B) of students with a common set of items that appear in each test. The two tests make it possible to compare the performance of one group of students to another group of students on the same set of items (College Entrance Examination Board, 2001). The Rasch model is then used to calibrate the items. For example, let us assume we have 98 items that need to be equated as shown in Figure 2. First impression would certainly indicate the number of items (e.g. 98 items) is far too large for a one group of students to complete. In this case, two tests are created for two groups of students. Test A is comprised of items 1 to 54 and Test B is comprised of items 45 to 98. This means 10 common items appear in each test.

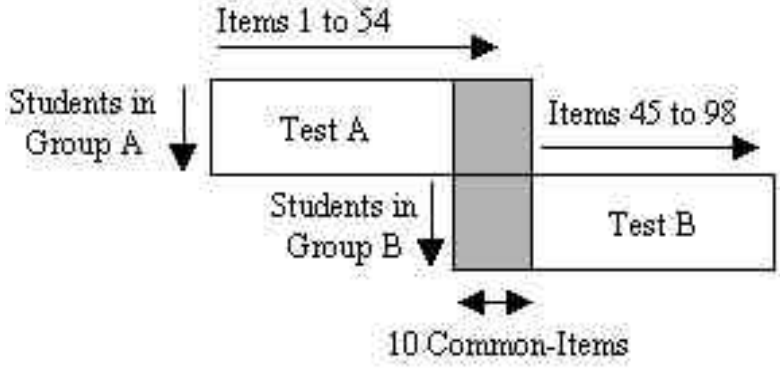


Figure 2. Common-Item Equating of two tests

Item Calibration using ConQuest

The raw data collected from the two tests (e.g. common-item equating) is then entered into a computer program called *ConQuest* (Wu et al. 1997). *ConQuest* is used to estimate the relative difficulty values of items from the raw data on the log-linear scale (Rasch Scale). It also provides analysts with up to date "psychometric methods of multifaceted item response models, multidimensional item response models, latent regression models and drawing plausible values" (see Australian Council for Educational Research ACER, 1999). Missing responses from students in each test such as omitted responses and not reached responses were given a value of 0 (see Keeves and Alagumalai, 1999). The next stage is to establish a database for LOT and CAT.

ESTABLISHING A DATABASE FOR AN ITEM-BANK

Two databases are required for computerised testing, one for linear on-line and another for adaptive testing. A database refers to a collection of data that is stored in related files (Shelly et al. 1997). For example, a personal telephone record may include a persons first and surname, date of birth, home and mobile telephone number and address. In this paper,

Microsoft Access 2000 is referred to as the default database. The basic design of a CAT and LOT database was adapted from Alagumalai (1995, 1998, 2000). The database for LOT and CAT consisted of 152 multiple-choice items. Each database is similar in design, except LOT has an added feature, Blooms Taxonomy of Objectives (pskill). The results of item calibration are entered into CAT and LOT databases under the attribute (rindex). All items in the item-bank record have a difficulty scale, which generally falls within the range -4 to +4 logits on the log-linear scale, as mentioned previously. Items in the database allow for test administrators (teachers) to edit, update or delete questions in this record. However, any changes made to the item bank record should not violate the principles of Item Response Theory (IRT) and the Rasch Model.

For the purpose of this paper, only two databases such as CAT item bank and the other for LOT item bank are described below. For further information see Alagumalai (1995, 1998, 2000).

CAT Item-Bank Record

The first to be described is the CAT Item Bank (CAT_item_bank). In order to describe the contents in this table, it is necessary to divide the records into two parts, the first part is the CAT item record and the second is the graphical record. Details are discussed as follows.

Item bank record contains the item number (QNo), question type (Qtype), year and semester (Qyear_sem), subtopic (Qtopic), question statements (Stem, alt_A,…, alt_E), keys (key_A,…, key_E) and item difficulty level (rindex). Figure 3 shows the basic structure of the item bank record for computer adaptive tests in Microsoft Access 2000.

Qno	Qtype	Qyear_sem	Qtopic	Stem	Alt_A	Key_A

alt_B	Key_B	alt_C	Key_C	Alt_D	Key_D	Alt_E

key_E	Rindex

Figure 3. The basic structure of item bank record for Computer adaptive tests

Item bank graphical record follows after the last file above (rindex) and contains the graphic indicator format yes/no (gf1_img,…, gf6_img), file graphic saved as (gf1_name1…, gf6_name6) and graphic description (gloc1…, gfloc6). Figure 4 shows the basic structure of the item bank graphic record for adaptive testing in Microsoft Access 2000.

LOT Item-Bank Record

The second to be described is the LOT Item Bank (LOT_item_bank). In order to describe the contents in this table, it is necessary to divide the records into two parts, the first part is the LOT item record and the second is the graphical record. The LOT graphical record is the same as Figure 4.

gf1_img	Gfname1	gfloc1	gf2_img	Gfname2	Gfloc2	gf3_img

gfname3	Gfloc3	gf4_img	Gfname4	Gfloc4	Gf5_img	Gfname5

gfloc5	gf6_img	Gfname6	Gfloc6

Figure 4. The basic structure of an item bank graphical record for CAT and LOT (From Alagumalai (1995, 1998, 2000).

Item bank record contains the item number (Index), year level (level), subtopic (stopic), item type (itype), Blooms Taxonomy (pskill), question statements (Stem, alt_A,…, alt_E), keys (key_A,…, key_E) and item difficulty level (rindex). Figure 5 shows the basic structure of the item bank record in Microsoft Access 2000.

index	Level	Stopic	Itype	pskill	stem	alt_A

key_A	Alt_B	key_B	Alt_C	key_C	alt_D	key_D

alt_E	Key_E	Rindex

Figure 5. The basic structure of item bank record for Linear online tests

SUMMARY

When assembling an item-bank, considerations need to be made for planning and assembling items, calibrating and entering items into a database before computerised testing can commence. These processes take much time and effort. They require patience and computer expertise. This paper was limited in the amount of information given on assembling an item-bank for computerised testing in Geography (see Burghof, 2001). However, there is a Geography item-bank ready and waiting for teachers to use, to edit items and populate the bank even further. It is also hoped that future item-banks are assembled for various other curriculum areas to be used in schools.

REFERENCES

Alagumalai, S. 1995, 1997, 1998. *ADAPTEST: Adaptive Testing System*, Trident Group, Singapore.

Alagumalai, S. 2000. *LOTS: Linear Online Testing System*, Wellington, New Zealand.

Alagumalai, S. 2000. *CATS: Computerised Adaptive Testing System*, Wellington, New Zealand.

Australian Council for Educational Research (ACER), 1999. 'ConQuest: Generalized Item Response Modelling Software' *Rasch Measurement Transaction*, Vol. 13, No. 3, p. 711 [Online] Available at http://www.rasch.org/rmt/rmt133o.htm [2001, September 22]

Burghof, K.L. 2001. *Computerised Adaptive Testing (CAT) in Society and Environment (SOSE): An exploratory study in Year 8/9 Geography, maps and mapping skills*, Unpublished B Ed Honours Thesis, School of Education, Flinders University, Adelaide

College Entrance Examination Board. 2001. 'Advanced Placement program: Common-Item Equating' [Online] Available at http://www.collegeboard.org/ap/techman/chap3/common.htm [2001, October 26]

Frary, R.B. 1995. More multiple-choice item writing do's and don'ts. *Practical Assessment, Research & Evaluation*, Vol. 4, No. 11. [Online] Available at http://ericae.net/pare/getvn.asp?v=4&n=11 [2001, September 22]

Granello, D. 2001. 'Promoting cognitive complexity in graduate written work: Using bloom's taxonomy as a pedagogical tool to improve literature reviews', *Counselor Education and Supervision*, Vol. 40, No. 4, pp. 292-308.

Keeves, J.P., and Alagumalai, S. 1999. 'New Approaches to Measurement'. In Masters, G. N., and J. P. Keeves (Eds), *Advances in Measurement in Educational Research and Assessment*, Pergamon, pp. 23-48.

Marsh, C. 1998. *Teaching Studies of Society and Environment* (2nd Edn), Prentice Hall, New South Wales, Australia.

McMillan, J.H. 1997. *Classroom Assessment: Principles and Practice for Effective Instruction*, Allyn and Bacon, Needham Heights, MA.

Microsoft Access 2000 (Windows 95/98) [Computer Program] Microsoft Corporation, Available Distributor: Computer stores

Microsoft Image Composer Version 1.5 (Windows 95/98) [Computer Program] Microsoft Corporation, Available Distributor: Computer stores

Microsoft Paint (Windows 95/98) [Computer Program] Microsoft Corporation, Available Distributor: Computer stores

Rudner, L.M. 1998a. 'An On-line, Interactive, Computer Adaptive Testing Tutorial', *ERIC Clearinghouse on Assessment and Evaluation* [Online] Available at http://ericae.net/scripts/cat/catdemo.htm#whatwhy [2001, April 05]

Rudner, L.M. 1998b. Item banking. *Practical Assessment, Research & Evaluation*, Vol. 6, No. 4. [Online] Available at http://ericae.net/pare/getvn.asp?v=6&n=4 [2001, September 22]

Sands, W., Waters, B., and McBride, J. 1997. *Computerized Adaptive Testing: From Inquiry to Operation*, American Psychological Association, Washington, DC, USA.

Theobald, J. 1974. *Classroom Testing: Principles and Practices* (2nd Edn), Longman Cheshire, Melbourne, Australia.

TOEFL, 2001. 'About The Computer-Based TOEFL Test', torfl.org [Online] Available at http://www.toefl.org/cbabttfl.html [2001, March 26]

Umar, J. 1999. 'Item Banking', In Masters, G. N., and J. P. Keeves (Eds), *Advances in Measurement in Educational Research and Assessment*, Pergamon, pp. 207-219.

Usova, G. 1997. 'Effective test item discrimination using Bloom's taxonomy', *Education*, Vol. 118, No. 1, pp. 100-110.

Wood, R., and Skurnik, L.S., 1969. *Item Banking*, National Foundation for Educational Research in England and Wales, Slough.

Wright, B.D., and Stone, M.K., 1979. *Best test design*, Mesa Press, Chicago.

Wu, M., Adams, R.J. and Wilson, M.R. (1997). *ConQuest: Generalised Item Response Modelling Software*, ACER, Melbourne.

Digital literacies and teachers in the middle years: Concepts and practices

Garth Coulter
Flinders University, School of Education *foster04@ihug.com.au*

This paper gauges teachers' views about the essential computing skills and understandings that students in years 6 – 9 require, and ascertains how these skills are supported and developed in teaching practices. Middle years teachers from five schools, located in country and metropolitan areas, participated in the study by completing four research instruments – online survey, questionnaire, classroom observation and interview. Through triangulation, the study seeks to establish the digital literacy competency levels of teachers. A key finding is that except for one area (word processing), digital literacy skills, as rated by the teachers themselves, fall in the two lowest competency bands of four competency levels. The findings from this research provide a foundation for professional development programmes and curricular objectives for teachers in the middle school.

Digital Literacies, Teachers, Competencies, ICT, Professional Development

INTRODUCTION

We are living in a society that has 'gone digital'. Mobile phones that comprised analogue signals have become a digital modernisation. Analogue VHS videos at the store have now succumbed to being replaced by Digital Video Discs (DVD). Television is now following suit into the digital revolution that is dominating today's society. Computers have become a necessary part of this digital society, and skills for computer use are a common prerequisite on many job applications, as the Department of Education, Training, and Employment (2001a) states:

> To live and work in the technology-enabled world of the 21st Century, high-level skills in the use of information and communication technologies (ICT) are essential for all citizens. (DETE, 2001a, p.5)

Information and communications technology has developed into being an imperative part of student learning at schools. New terms such as 'digital' and 'online' information have been embedded in schools' curriculum across the world so that future society will continue to support and progressively develop a digital world.

In light of the continual development of students' skills in computers and technology, the term digital literacy seems to describe best the skills, knowledge, and understanding of computer uses and practices required in today's society. Gilster (1997) presents a useful definition of digital literacy that is adopted in this paper:

> Digital literacy is the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers. (Gilster, 1997, p. 1)

The use of digital literacy constitutes an educational concern for facilitating teachers, as the development of digital literacies, for students in the middle years, require teacher computer knowledge, computer use and digital literacy understanding. The curriculum criteria for

students in schools are continually progressing with the use of digital technologies, as stated by DETE (2001a):

> All students will leave school as confident, creative and productive uses of new technologies, particularly information and communication technologies and all students to 16 years will have received in education in the use of information technology telecommunications, and will leave school with an industry-recognised information technology qualification. (Directions for South Australia 2000-01, Government of South Australia; cited in DETE, 2001a, p.5)

To support the ideals of digital literacies in the classroom, teacher views of what are necessary digital literacies for students in the middle years are defined, and the teaching practices involved in supporting and developing the necessary digital literacies are investigated. The present study focuses on the middle years of schooling for the reason that information technology has become a key subject integration in these years. Digital literacies are a significant area of study as the students in these years are susceptible to new information and practices.

AIMS AND PURPOSES

The primary aim of this study is to ascertain what teachers believe are the necessary digital literacies required for students in the middle years of schooling (years 6 – 9) and to find out what teaching practices support these necessary literacies.

In order to develop an understanding of teaching practices in the middle years, the digital literacy competencies of teachers in the middle years are viewed as an integral part of the teaching practices of digital literacies.

SIGNIFICANCE OF THE STUDY

Society's increasing demand for teachers to incorporate computer skills into their subject areas forces teachers to develop their own skills in computer use and learning (Meredyth et al., 1999). It is therefore important to determine the skills that teachers require in today's age of technology use. Previous studies reported that teachers lack computer literacy skills (Meredyth et al., 1999; Taylor, 1999), and that teacher professional development is required. No study has been sighted that specifically targets the support and development of digital literacies or that seeks to discover teacher views on what they perceive to be essential digital literacy skills for students in the middle years of schooling. This study is potentially important as it provides an active approach to the foundation of teacher professional development in the middle years. Establishing knowledge of the concepts and practices of teachers will assist the technology development of such programmes.

RESEARCH QUESTIONS

To study digital literacy concepts and practices of teachers in the middle years, three key research questions need to be considered:

1. What do teachers of the middle years view as necessary digital literacies?
2. What is the level of computer use and knowledge of teachers in the middle years, as these relate to digital literacies?
3. How is teaching practice in the middle years being used to support and develop digital literacies?

DEVELOPMENT OF RESEARCH INSTRUMENTS

In order to investigate the research questions stated above, four data collection instruments were developed. The first instrument is a questionnaire that was administered to a small sample of middle years teachers. The second instrument is an online test that was administered through the Internet to the same number of teachers in the middle years. The third instrument is an observation schedule of classroom practices in a particular middle school. The fourth instrument is an interview schedule designed for teachers.

Selection of Subjects

The study included both government and independent schools in country and city areas of South Australia. For reasons of convenience, three government schools (one primary and two secondary) and two independent schools (one years 6 – 9 and one primary) were included in the sample. Within each of these schools, teachers were approached to participate in the questionnaires, online surveys and the interviews/observation of teaching practices. There were 11 participants who completed the teacher questionnaire from independent schools (4 males and 7 females) and 45 participants from government schools (18 males and 27 females), making a total of 58 participants (23 males, 35 females).

The online teacher survey involved 58 teachers, 24 males and 34 females. There were 29 teachers from government schools (11 males and 18 females) and 27 from independent schools (12 males and 15 females).

One school was involved in the classroom observations, with eight classrooms being observed from Years 6-9 (two from each year level). Two male teachers and two female teachers participated in the teacher interview.

Teacher Questionnaire

A questionnaire was developed to address the first research question of what middle year teachers viewed as necessary digital literacies for their students. The process of developing this questionnaire involved revisiting the first research question. In order to discover what teachers view as necessary digital literacies, the questionnaire was designed in two parts. The first part of the questionnaire contained items which asked teachers what subject areas and over what year levels they practice. By including subject areas and year level questions, the factors of subject specifics and age range are considered. The items in the second part of the questionnaire asked what teachers believed to be necessary computing skills for students from Years 6-9, and to state their reasons why. This questionnaire directly addressed the first research question. However, by asking teachers to give reasons, teachers provide a qualitative response on why they view a particular skill as necessary for students from Years 6-9. In the design of the questionnaire, the process is to find a common overlap of digital literacies in order to develop a list of necessary skills for students and teachers in the middle years. The thinking behind the statement of subject areas assists in explaining the skill choices of teachers with key teaching areas.

The results from this questionnaire are summarized in Table 5.1, with the frequency of essential literacies listed by teachers.

Online Teacher Survey

The online survey is a quantitative instrument designed to answer the second research question concerning the current level of digital literacy of teachers. The questionnaire uses Likert-type items and is adapted from a measure used by Bellingham Public Schools (2001). The process of developing this survey was to revisit the second research question, which asks what level of computer use and knowledge a teacher has in the middle years.

Table 5.1 Frequency of Essential Digital Literacies as Listed by 56 Middle Years Teachers

Essential Digital Literacies	Frequency
Word Processing (Fonts, Font Sizes, Column/Margins, Pictures, WordArt, Clipart, Tables, Headers/Footers, Spell Check, Cut/Copy and Paste Skills	69
Internet (Searching, Viruses, Research, Evaluation, Library, Catalogue, Downloading, Chat Pages, Online Tutorials)	65
Computer Basics (Multitasking, Menu Bar, Navigation, Save and/or Name Files/Folders, Hardware Components, Renaming/Deleting/Copying Files and Folders, Operating System Skills, General Troubleshooting Skills)	39
E-mail	36
Spreadsheets (Graphs/Charts, Tables, Sorting, Formulas)	34
Graphics/Design (Photoshop Software, Basic Drawing Skills, FreeHand Software, Web Design, Geometric Patterns, CAD Software, Inspiration Software)	29
CD-ROMs	21
Publishing Programs (Layers/Frames, Graphics, Text Wrapping)	20
Presentation Software (PowerPoint)	11
Databases	9
Typing/Keyboarding	8
Scanners	5
Others	5
Video Editing Software	3
Digital Cameras	3
Audio Programs	3
Games	3

A list of 14 questions covering 4 different skill areas was designed with 4 multiple choice items under each question. Each of these four multiple choice items consisted of different measures of computer use and skill levels, ranging from not using a computer at all, to achieving a high band of competence and knowledge in computing. The skill levels themselves were adapted directly from the Bellingham Public Schools (2001) website so that a structured test of valid computing skills could be used.

The test items under each skill were then modified to match the different skill bands. The base band was adjusted to reflect if teachers did not use a particular skill at all, or whether they had a very basic knowledge of the skill. The next band was designed to gauge a basic knowledge that reflected if a skill was used by the teacher outside the classroom. The third band was redeveloped to reflect a higher level of confidence and ability with computing skills that allowed the teacher to competently use the computer outside the classroom. The final band was redesigned to reflect the highest based skill being able to competently use and teach information and communication technology skills in the classroom.

The readjustment of skill bands in the Bellingham Schools form (2001) was necessary in order to allow greater distinction between each band, and to clarify results relating to the second research question.

The decision to use an online survey allows a relatively large number of teachers to be involved in the research study in different areas.

The results from this survey are clearly depicted in Figure 5.1.

Figure Not Available

Figure 5.1 Means for 56 Teachers to 14 Items in the Online Survey

Examination of Figure 5.1 clearly shows means for 56 teachers in relation to the four competency bands, highlighted by separate colour levels, in the online instrument. For example, in only one area, namely word processing, do teachers score in band 3 ($x = 3.11$), that is, they are able to confidently use word processing skills in a classroom situation and are able to construct personal documents easily. Figure 5.1 also shows that for four areas of ICT (databases, desktop publishing, video presentation software, and presentation software), teachers rate themselves at a competency below band 2. In other words, teachers have limited knowledge of the use of computers in a teaching situation, but they can use pre-loaded programs for basic purposes. For database applications, the mean is 1.79, for desktop publishing the mean is 1.88, for video production it is 1.63, and for technology presentation it is 1.91.

Classroom Observation

The decision to observe digital literacy learning in classroom settings allows a comprehensive analysis of the third research question. By observing how teaching practices are being used to support the different digital literacies, the results from the first test instrument were used as a checklist of what skills were being developed in the classroom. It also highlighted the student learning factors and teacher's own knowledge base involved in teaching. The process of developing classroom observation checklists involved an analysis of the results obtained from the questionnaire instrument. This checklist was developed by setting up a spreadsheet and designing skill level headings under each specific subject area.

Classroom observations were mapped to cover Years 6-9, observing two classes at each year level, incorporating a range of subject areas in a single gender middle school. In order to comply with ethical standards, the School Principal was approached for permission as well as the classroom teachers. Each classroom teacher signed a form permitting the observation study in his or her classroom.

The observed teaching practices proved to clearly support digital literacies in these classrooms, and a range of computing skills were covered when viewing the necessary digital literacies posed earlier in the chapter.

Teacher Interview

The teacher interview was designed to cover all three research questions, providing an overlap with the previous test instruments. The interview was developed with 10 questions that addressed computer use and application in classrooms, covering questions asked in the first and second instruments. Four teachers (two male and two female) were interviewed in a single school.

The development of the questions was purposefully designed in order not to present leading questions or criteria that did not specifically relate to the research questions. The interview was designed to be conducted within a time frame of 30 minutes, allowing sufficient time to interview practicing teachers between classroom periods.

The interviews were recorded on tape and then transcribed for analysis.

The literacies that teachers used were a part of the interview. These digital literacies are shown in Table 5.2

CONCLUSIONS

The main purposes of this study were to determine what teachers deemed to be essential digital literacies for students in the middle years and how teachers supported and developed these digital literacies. Results were collected from questionnaires and interviews. A further objective was to measure teacher skill competencies in computer use and application in the

classroom, related to digital literacies. Results for this objective were collected through an online survey.

Table 5.2 Computer Digital Literacies Used in Subjects by Four Teachers

Subjects	Digital Literacies	
English	PowerPoint	
	E-mail	
	Internet	
	Word Processing	(Clipart, Fonts, Paragraphing, Spell Checking, Columns & Margins)
Religion	PowerPoint	
	Desktop Publishing	
	Internet	(Pictures, Online News)
Visual Arts	Graphic Programs	
	Desktop Publishing	
	Word Processing	
Society and Environment	CD-ROMs	
Maths	Spreadsheets	(Inserting and using Formulas, Graphs)
Science	CD-ROMs	
	Internet	(Locating and Inserting Graphics, References)

1. Teachers rated Word processing, Internet Skills, Computing Basics, E-mail, Spreadsheets, Graphics/Design, CD-ROMs, and Publishing Programs the most essential digital literacies for students in the middle years. The less frequent essential literacies teachers listed were Presentation software, Databases, Typing/Keyboarding Skills, Scanners, Video Editing Software, Digital Cameras, Audio Programmes and Games.

2. Teachers rated themselves as most competent in the area of word processing. The average performance rating in the majority of competencies was located within the third band of competency. Teachers ranked themselves most lowly in Databases, Desktop Publishing, Video Production, and Technology Presentation, which fell below the second band of competency. The remaining teacher digital literacies were Computer Use, File Management, Spreadsheets, Graphics, E-mail, Research/Information Searching, Internet, Responsible Use/Ethics and Embedding Technology, which the average of the 56 teachers rated themselves within the second band of competency.

3. Data collected from classroom observations showed that teaching practices supported student digital literacies in the classroom. However, the teacher interviews revealed that, on average, there is little evidence of digital literacy development in classrooms through the lack of skill assessment and curriculum structure.

A key feature of the research design employed in this study was the collection of data using

multiple perspectives – through questionnaire, online survey, classroom observation, and teacher interview. This triangulation approach helps increase the validity of the findings, according to many writers (Jacob, 1990; O'Malley and Valdez Pierce, 1996; Maxwell, 1996; Wiggins, 1998). As a result, one can feel more confident in the research findings, even though these findings are based on relatively small numbers.

The results clearly indicate which digital literacies teachers can improve on and how they can be further developed for students in the classroom. The data results from the second research show an average of teacher competency with computer knowledge and use. It is important to

note, however, that the findings of computer competency are based on the averages of 56 teacher responses. They do not show individual teacher competency. Individually, there were teachers who rated themselves in the fourth band of competency. Table 6.1 shows the number of individual teachers who rate themselves in the fourth band of competency. The fourth band of competency is the highest band that generally shows teachers have a firm grasp on particular digital literacy concepts and uses.

Further studies in digital literacies are necessary for the progress of education and its use in society. There is great potential for future research in areas of professional development for teachers and parents, and further developments of student learning in digital literacies. With continuing research, it seems highly likely that schools will adopt an embedded ICT curriculum with in service training for teachers in the middle years, resulting in the continually increasing teacher knowledge and understandings of digital literacy concepts and practices.

REFERENCES

Bellingham Public Schools. 2001. Staff Use of Technology. [Online] Available: http://www.bham.wednet.edu/StafSlfAsmt.htm [Accessed: October 2001]

Department of Education, Training and Employment. 2001a. Certificate 1 in Information Technology, DETE Publishing, South Australia.

Department of Education, Training and Employment. 2001b. Learning Futures: DECStech 2001, DETE, South Australia.

Gilster, P. 1997. Digital Literacy, John Wiley & Sons, Canada.

Jacob, E. 1990. Alternative approaches for studying naturally occurring human behavior and thought in special education research. The Journal of Special Education, 24,(2), 195-211.

Maxwell, J. 1996. Using qualitative research to develop causal explanations, Harvard University: Harvard Project on Schooling and Children.

Meredyth, D., Russell, N., Blackwood, L., Thomas, J. and Wise, P. 1998. Real Time: Computers, Change and Schooling. Canberra, Department of Education, Training and Youth Affairs.

O'Malley, J., Valdez P. 1996. Authentic assessment for english language learners: Practical approaches for teachers. Reading, MA: Addison-Wesley Publishing Company.

Taylor, P. 1999. IT Skills and Schooling: Rethinking Assumptions about Access, Griffith Institute for Higher Education, Griffith University. [Online] Available: http://www.aare.edu.au/99pap/tay99146.htm [Accessed: April 2001]

Wiggins, G. 1998. Educative Assessment. San Francisco, CA: Jossey-Bass.

Misfits: People and their problems. What might it all mean?

David D Curtis
Flinders University, Centre for Lifelong Learning and Development and School of Education
david.curtis@flinders.edu.au

In the analysis of data, which arise from the administration of multiple choice tests or survey instruments and which are assumed to conform to a measurement model such as Rasch, it is normal practice to check item fit statistics in order to ensure that the items used in the instrument cohere to form a unidimensional trait measure. However, checking whether individuals also fit the measurement model appears to be less common. It is shown that poor person-fit compromises item parameter estimates and so it is argued that person-fit should be checked routinely in the calibration of instruments and in scoring individuals. Unfortunately, the meanings that can be ascribed to person-fit statistics for attitude instruments is not clear. A proposal for seeking the required clarity is developed.

Item Response Theory, Rasch, person-fit statistics, attitude

Three sets of data derived from the application of different attitude survey instruments have been analysed using Item Response Theory (IRT) based software packages Quest (Adams & Khoo, 1997) and Conquest (Wu, Adams, & Wilson, 1998). In all cases, items that fit a unidimensional scale have been found. Kline (1993) however, has noted that the Rasch (one parameter) IRT model may be insensitive to departures from the assumption of unidimensionality. However, closer examination of the data using Confirmatory Factor Analysis (CFA) has found that the structure, supported in IRT analysis, is regarded as only moderately fitting and that further refinement was required. A fundamental requirement of measurement is that the set of items represent a single, unidimensional construct (Michell, 1997; Weiss & Yoes, 1991; Wright & Masters, 1982).

In a review of these data sets, a significant number of poorly fitting cases were found. These cases may have compromised the calibration of the instrument and may have introduced factors other than the attitude construct, whose measurement was the goal of the instruments. The IRT software appears to have been less sensitive to sources of variation in the data than the CFA software. In IRT based measurement, the claim is made that item parameters are independent of the particular sample of respondents and that respondent scores are independent of the particular sample of items used to measure the underlying trait. Thus both items and persons are samples of all possible items and of the universe of persons represented by the sample. But are misfitting persons somehow not legitimate members of a coherent population? What criteria might be used to disqualify them?

In seeking answers to these questions, the person-fit statistics were reviewed. It seems possible that misfitting persons may introduce a source of variance that is detected through CFA but that is not detected using IRT software. If this is the case, it is possible that item calibrations are suspect and even that acceptable items have been rejected in the IRT refinement of the instrument. In addition, attitude survey instruments are the central concern of this research, and criteria for judging the appropriateness of response patterns are rather different from criteria that might be used in knowledge testing. In objective tests of knowledge, cheating and

guessing are threats to valid response patterns. However, in attitude surveys these are not viable threats, although others, such as a desire to appear to have favourable views, may be.

In order to investigate the effect of poorly fitting cases on item parameter estimates, data from the 1996 application of the Course Experience Questionnaire (CEQ) were reanalysed.

A REVIEW OF PERSON-FIT STATISTICS

Fit statistics for both items and persons are derived from deviations of observed responses from expected ones. The matrix of individuals' responses to items is tabulated and the rows and columns summed. Under the Rasch model, the person and item totals are used to estimate person ability (β) and item difficulty (δ). In turn, these values are used to estimate the expected response of each person to each item. In general, an observed response should be associated with a high probability of that response, and the deviation of the observed response from the expected one is an indicator of misfit.. The simplest misfit statistic is the unweighted mean square residual, found by taking the mean value of the squared differences between the observed and expected responses, and is the Outfit Mean Square reported in Quest.

Karabatsos (2000) is critical of fit statistics such as this, arguing that the expected value is a continuous quantity while the observed response is always a discrete one – most often a '1' or a '0'. This means that, even for very well fitting items and persons, there will always be a residual, albeit a small one. This may not be important, as fit is a matter of relative judgement, and for poorly fitting items or persons, residuals will be much higher than for those with reasonable fit. Karabatsos has proposed an alternative approach to item and person misfit, that itself is based on a logisitc model, using the measurement scale of the instrument.

It appears to be common practice in the analysis of data sets under IRT, collected for purposes of validating instruments, to focus on item parameters and to pay little attention to case fit statistics (eg Waugh, 1999). It might be added that other methods of analysis do not attend to the issue of person-fit at all and that the sampling process is assumed to generate a genuinely representative set of individuals. In some of the original treatments of the Rasch Measurement Model (eg Wright & Masters, 1982), almost equal treatment is accorded to both item and person-fit statistics. Indeed, there appears to be a sound case for such equity. Inspection of the matrix of any data set being analysed using the Rasch model reveals that analogous operations are performed on the rows and columns to derive marginal totals and expected cell values. Whether the primary view is of items or of people, the mathematics of the analysis for items and persons are mirror images, and if the matrix is to be fitted to a measurement model, then both the items and the persons should be treated similarly.

A range of person-fit statistics has been used to assess conformity of the data set to the measurement model (Li & Olejnik, 1997). They tested five indices of misfit, including those that are employed in the Quest and Conquest programs – the Weighted Mean Square residuals (WMS or Infit Mean Square) and the Unweighted Mean Square residuals (UMS or Outfit Mean Square) developed by Wright and Stone (1979). Although Li and Olejnik reported that there was little difference among the performances of the indices that were tested, the performance of the two that are used in Quest was slightly inferior in both the detection of misfit and in reporting false positives, that is reporting as misfits cases that were reasonable fits. They further reported that these indices do not strictly follow a normal distribution, and recommend that their standardised versions, Infit t and Outfit t, should be given somewhat more latitude than the common ± 2.

In the current study, there are many cases (51,356) and relatively few items – 17 after refinement of the instrument. This means that item estimates are based on many cases, and very small misfits have very high Infit t values. Case estimates, based on relatively few items, can show substantial misfit but still have modest Infit t values. This raises the question: Should fit be judged on the Infit MS or the Infit t value? In addition, in large samples like the one

being reported here, there is a degree of clustering within course types and within institutions, and any analysis that treats all cases as being drawn from a single population will reveal aggregation bias and consequently false significance.

The study reported by Li and Olejnik (1997) simulated objective test data. In such tests, there are a range of threats to measurement model conformity. Most commonly, carelessness, cheating, lucky guessing, special knowledge, and miscoding of the data are proposed as possible reasons for atypical response patterns (Bond & Fox, 2001, p.178). These conditions normally lead to high (under-fitting) person-fit statistics. They also indicate that over-fitting can arise when the responses follow a deterministic response pattern whose fit statistics are "too low to be believed" (Bond & Fox, 2001, p.178). The reasons advanced to explain poor person-fit are relevant to objective testing, but not necessarily to attitude survey instruments. In these cases cheating or guessing seem irrelevant, but carelessness or even deliberate spoiling of the instrument and miscoding may be problems.

In the current study, attitude data derived from a survey have been re-analysed in order to examine possible sources of misfit and to ascertain the effect of person misfit on item parameter estimation.

METHOD

Data from the 1996 round of the Course Experience Questionnaire (CEQ) was analysed previously under IRT using Quest (Adams & Khoo, 1997). The CEQ is a 25 item instrument designed to gather the views of graduates on the quality of their recently completed courses. Respondents selected one of five options – Strongly Disagree, Disagree, Neutral, Agree, or Strongly Agree to each item. The CEQ was based on a theoretically informed view of the major contributing factors to quality of teaching and to perceived course quality and comprised five sub-scales. The refinement of the instrument revealed that only 17 of the items fitted both their intended sub-scale and an overall coherent scale of perceived course quality. Subsequent CFA revealed that a nested structure of one underlying course quality construct and the five proposed sub-scale factors was a reasonable model (Curtis, 1999, p.18).

In the current study, the 17 fitting items were taken as a starting point and Rasch scaled scores, standard errors, and fit statistics were generated for the 51361 cases for whom complete data were available. Of them, 65 (0.13 per cent) had either zero or perfect scores, 7093 (13.74 per cent) had low Infit MS (<-2), and 5384 (10.43 per cent) had high Infit MS values (>2).

In order to build a case for the exclusion of persons who demonstrated misfitting response patterns, the responses of the most poorly fitting cases, both those with very low and very high Infit MS values, were examined. The cases with zero or perfect scores are not included in item or case estimation as the algorithm cannot estimate scaled case scores nor produce fit statistics for them. These are people who have responded consistently to items in that they have either very positive sentiments (agreed strongly with positive items and disagreed strongly with negative items) or they have very negative views (having disagreed strongly with all positive items and agreed strongly with all negative items). There are methods for extrapolating from scaled cases and imputing scaled scores for these individuals, although no confidence interval can be associated with these imputed values. These cases are not of interest in the current study as they are excluded from Rasch analysis and do not influence item or person parameter estimation.

Inspection of the response patterns of 20 cases showing the highest misfits revealed that respondents had selected either 'Strongly Disagree' or 'Strongly Agree' to all items, even though eight of the items were reversed. Thus their responses were inconsistent with a coherent expression of the trait targeted in the instrument. The pattern of these cases is that they have checked all responses down either the left or the hand side of the response column, irrespective of the sense (positive or negative) of the items. Both the item fit statistics and the

observed pattern of responses indicate that there is a reasonable case for removing the responses of these people.

The response patterns of the cases with the lowest person-fit statistics (with Infit MS values around 0.10) were less simple. These cases are characterised by the choice of middle range response options, either 'Agree' or 'Neutral', but more importantly by an invariant response pattern. Typically these respondents selected 'Agree' to all items but one, and selected an adjacent category, say 'Neutral' to the remaining item. Item locations vary and a person with a true level on the underlying trait of perceived course quality that falls close to the range represented by this response category ('Agree') might be expected to select an adjacent category for some of the items. However, selecting the 'Agree' response option is not an unlikely event for a person who has a moderate overall trait score. Therefore the case for removing these persons from the analysis is far from convincing. Thus, for persons with low Infit MS values, the person-fit statistics might suggest removal of the cases, at least for calibration purposes. However, the pattern of responses cannot be interpreted as easily as in the situation for under-fitting cases described above. On one hand, it could be argued that the invariant selection of a single response option suggests a 'patterned and thoughtless response' but the responses are not improbable given the underlying trait value. This analysis suggests that other criteria require examination before a decision is made to retain or to exclude over-fitting cases.

In order to ascertain the influence of misfitting cases on item parameter estimates, the original data set with all 51,631 cases was used to generate item parameters. Subsequently, under-fitting cases (Infit $t > 2$) were removed and item parameters re-estimated. Finally, all misfitting cases (Infit $t < -2$ and Infit $t > 2$) were removed and item parameters again re-estimated. In order to establish further the influence of person-fit on overall model fit, confirmatory factor analyses were undertaken. A refinement process using the Quest software was also conducted beginning with all 25 items and successively removing misfitting items with both the under-fitting cases removed and then with all misfitting cases excluded. This was done to examine whether misfitting cases would bias the fit statistics of items and therefore the items removed during the refinement process.

RESULTS

The Influence of Person-fit on Item Locations

While for some items, for example Item 1, the effect of removing misfitting cases has little impact, for the majority the effect is significant. This is shown in ***Table 1***. Here, Item 1 shows a small change, but Item 5 shows a shift of 0.13 logits when under-fitting cases are removed from the calibration. This is taken to be a significant effect given that the standard errors of the location estimates are 0.01 for all items. In summary, the removal of misfitting cases leads to a greater spread of item locations. This effect is more pronounced for under-fitting cases and suggests that misfitting cases add to error variance but not to information about items.

The Influence of Person-fit on Item Thresholds

The influence of cases with poor person-fit statistics were shown also to influence item threshold parameters. The four threshold parameters for each of the 17 items with all cases included, under-fits removed, and all misfits removed are shown in ***Table 2***. Standard errors for the threshold parameter estimates, which are not shown in the table, vary from 0.01 to 0.06, with most at 0.02 logits. The thresholds under the three conditions vary by considerably more than could be attributed to random error. In general, the inclusion of misfitting cases, especially under-fitting ones, compresses the width of the item steps on average by 0.32 logits with under-fitting cases are removed and by 0.25 logits with all misfitting cases are removed.

Of particular note are the locations of the first and fourth thresholds. The inclusion of all cases results in a compression of the thresholds compared with the situation in which the under-fitting cases are excluded. This is

illustrated most effectively by tabulating differences between the first and fourth thresholds under the three conditions and is shown in

Table 3. The removal of under-fitting cases extends the range by almost a logit while also removing over-fitting cases extends the range by approximately one half of a logit compared with the situation in which all cases are retained. This has implications for later applications of a scale. Calibrating the scale on one sample and anchoring item parameters, so that case scores may then be estimated for other samples, may lead to better person separation in subsequent samples.

Table 1: Item locations for three person-fit conditions

	Item Locations		
Item No.	All cases	Under-fits removed	All misfits removed
1	0.06	0.07	0.06
2	-0.46	-0.58	-0.55
5	-0.53	-0.66	-0.63
6	-0.03	-0.05	-0.05
7	0.64	0.77	0.73
10	-0.12	-0.19	-0.18
11	-0.37	-0.44	-0.41
12	-0.09	-0.06	-0.05
13	0.06	0.05	0.05
14	0.17	0.25	0.24
15	0.37	0.46	0.44
17	0.41	0.50	0.48
18	0.32	0.39	0.37
19	-0.42	-0.48	-0.47
20	0.17	0.21	0.19
22	-0.42	-0.52	-0.50
24	0.24	0.29	0.27

Standard Errors are not shown, but for all items, the standard error of the estimate is 0.01 logits.

Confirmatory Factor Analyses

In the original analyses of the 1996 CEQ data, confirmatory factor analyses were undertaken in order to identify the most suitable model for representing the structure of the instrument. These analyses suggested that a nested model was the most appropriate (Curtis, 1999). In the present study, the nested model was taken as established, and the model was re-run under LISREL 8.12a (Joreskog & Sorbom, 1993) in order to examine the impact of the removal of misfitting cases on model fit. ***Table 4*** shows that the removal of misfitting cases, and most particularly the removal of under-fitting cases, results in a slight improvement in model fit. However, the parsimony fit index is rather low in all three conditions, suggesting that the model might be further simplified.

RASCH REFINEMENT OF ITEMS

In order to examine whether the removal of items in the original refinement was a result of the inclusion of misfitting cases, the refinement process was conducted after excluding under-fitting cases and again following the removal of all misfitting cases. With both reduced data sets, all 25 items were included and misfitting items were removed in an iterative process. For both reduced data sets, the same eight items that had been cut in the original refinement were

again removed. This result suggests that, despite having a sample of which 25 per cent of cases showed poor fit, item fit statistics are not strongly influenced by the presence of these cases and the refinement process is a reasonably robust one.

Table 2: Item thresholds for three person-fit conditions

	Tau 1			Tau 2			Tau 3			Tau 4		
Item No.	All cases	No under-fits	No misfits	All cases	No under-fits	No misfits	All cases	No under-fits	No misfits	All cases	No under-fits	No misfits
1	-1.85	-2.35	-2.19	-0.78	-0.86	-0.66	0.35	0.45	0.34	2.28	2.76	2.51
2	-1.46	-2.00	-1.87	-0.65	-0.68	-0.50	0.11	0.26	0.22	2.00	2.43	2.15
5	-1.40	-1.95	-1.81	-0.77	-0.82	-0.64	0.15	0.31	0.27	2.02	2.46	2.19
6	-1.58	-2.07	-1.90	-0.62	-0.67	-0.49	0.11	0.21	0.11	2.09	2.54	2.28
7	-1.60	-1.97	-1.73	-0.64	-0.74	-0.61	0.48	0.54	0.38	1.76	2.17	1.97
10	-1.53	-2.10	-1.94	-0.91	-0.95	-0.77	0.28	0.41	0.34	2.16	2.64	2.37
11	-1.05	-1.47	-1.33	-0.46	-0.50	-0.32	-0.01	0.09	0.05	1.52	1.88	1.60
12	-1.09	-1.43	-1.26	-0.63	-0.69	-0.51	0.31	0.37	0.27	1.40	1.76	1.50
13	-1.73	-2.26	-2.09	-0.72	-0.78	-0.59	0.37	0.48	0.38	2.08	2.56	2.30
14	-1.92	-2.43	-2.23	-0.82	-0.94	-0.75	0.24	0.30	0.16	2.51	3.07	2.82
15	-1.56	-2.01	-1.79	-0.75	-0.84	-0.69	0.48	0.56	0.41	1.84	2.30	2.07
17	-1.74	-2.19	-1.97	-0.70	-0.80	-0.63	0.40	0.46	0.30	2.05	2.53	2.30
18	-1.74	-2.24	-2.02	-1.10	-1.19	-1.03	0.65	0.74	0.60	2.19	2.68	2.45
19	-1.61	-2.22	-2.08	-1.29	-1.33	-1.13	0.83	0.97	0.90	2.07	2.59	2.31
20	-1.57	-2.04	-1.84	-0.97	-1.05	-0.89	0.38	0.47	0.34	2.16	2.62	2.38
22	-1.21	-1.80	-1.66	-0.78	-0.80	-0.58	0.09	0.23	0.16	1.90	2.36	2.08
24	-1.81	-2.34	-2.14	-0.84	-0.93	-0.74	0.53	0.63	0.49	2.12	2.64	2.39

Note that standard errors are not shown. They vary from 0.01 to 0.06 with 0.02 being the modal value.

Table 3: Item threshold range for three person-fit conditions

	Threshold range Tau 4 - Tau 1		
Item No.	All cases	Under-fits removed	All misfits removed
1	4.13	5.11	4.70
2	3.46	4.43	4.02
5	3.42	4.41	4.00
6	3.67	4.61	4.18
7	3.36	4.14	3.70
10	3.69	4.74	4.31
11	2.57	3.35	2.93
12	2.49	3.19	2.76
13	3.81	4.82	4.39
14	4.43	5.50	5.05
15	3.40	4.31	3.86
17	3.79	4.72	4.27
18	3.93	4.92	4.47
19	3.68	4.81	4.39
20	3.73	4.66	4.22
22	3.11	4.16	3.74
24	3.93	4.98	4.53

DISCUSSION

This study has shown that the inclusion of misfitting cases, especially under-fitting ones, during instrument calibration has a significant effect on item location parameters and on item threshold estimates. In particular, variation among item locations is reduced and item steps are truncated by retaining misfitting cases. However, poor person-fit does not seem to influence item fit statistics and the refinement of instruments to detect and remove misfitting items can proceed while retaining misfitting cases. Retaining poorly fitting cases does seem to compromise the overall model fit slightly when the hypothesised structure of the instrument is being evaluated using confirmatory factor analysis.

Table 4: Summary fit statistics from confirmatory factor analyses

	Nested factor model with all cases	Nested factor model with no under-fits	Nested factor model with no misfits
N	51631	46182	39089
χ^2/df	143.50	121.53	103.59
RMSEA	0.053	0.051	0.051
RMR	0.027	0.025	0.026
GFI	0.97	0.97	0.97
PMI	0.65	0.65	0.58

Note: RMSEA = Root Mean Square Error of Approximation; RMR = Root Mean Square Residual; GFI = Goodness of Fit Index; PMI = Parsimony Fit Index

The influence of person misfit on item parameters has implications for the calibration of attitude instruments and for routine procedures in checking model fit under the Rasch measurement model. Criteria for identifying person misfit in attitude surveys must be more firmly established, as these instruments cannot be taken as completely analogous to objective tests.

First, some of the causes of misfit in objective tests, such as guessing and cheating, do not apply to attitude surveys, but other threats to their measurement validity, such as carelessness or disinterest, may. Attitude surveys are not normally high stakes tests for the individuals taking them, and this is certainly true for the CEQ. The 51,631 cases analysed in this study represent a sub-set of all respondents. These were selected originally because they responded to all 25 items. Some 12,000 cases with incomplete forms had been removed previously. Many of those had only answered the first few items. Of the 5,384 under-fitting cases, it is not clear that all had followed a simple and thoughtless pattern of choosing only the first or the last option for all items. Closer analysis of patterns of responses is required before a sound rationale can be established for removing all under-fitting cases from data sets, and cut-off criteria that engender greater confidence than the simple 'Infit t > 2' rule that was applied in this analysis must be found.

Second, judgments of misfit are based on conditional probabilities of selecting particular response options to items for individuals with a given level on the latent trait. However, differences in the probabilities of selecting a particular response option compared with adjacent ones may not be great. This is partly attributable to the categories that are offered to participants in attitude survey instruments. What does 'Strongly Agree' mean compared with 'Agree' to a respondent, and does it have the same meaning for all respondents? Answers to these questions are far from clear, and this may account for much variance in respondents' choices of options. It is desirable to attach a much more precise meaning to each response option and to convey this clearly to respondents. It is instructive to compare this situation with

that of a rating scale in judging performance. Similar data are generated, but with judged rating scales, there are more exacting criteria to determine which option on the scale is to be applied. Judged rating scales are likely to yield more precise estimates of thresholds and of persons being rated than are attitude surveys unless greater clarity can be given to response options.

Third, in the Rasch measurement model, the claim is made that item parameters are (person) sample independent and that case estimates are item independent. However, it must be understood that the items used represent a universe of possible items that reflect the trait being assessed. Items that are judged not to 'fit the scale' are rejected as not being part of the universe of trait-related items: they reflect, or are contaminated by, a different trait. Similarly, persons who do not fit are judged not to conform to the typical response pattern of the majority of respondents. There may be several reasons for this. Some ill-fitting individuals may have responded carelessly, but there are other plausible reasons for poor person-fit. Some items may have particular salience for a subset of respondents and their responses to these items are likely not to conform with the 'standard' pattern. This is a case of differential item function (DIF). The inclusion of these persons in calibration may distort the parameter estimates for those items. However, with DIF there is always uncertainty about whether the item discriminates against some persons, that is whether there is item bias, or whether these persons have a genuinely different level of attitude (or performance) on that item. There is some evidence of DIF in the CEQ data as respondent sex, age, and NESB status were all found to be related to perceived course quality. If the structure of the instrument is to be validated, alternatives to simple confirmatory factor analysis are required. Where clear evidence of careless patterned responses is found, those cases can be removed. Where individual difference attributes are thought to be related to responses, such patterns need to be factored into the response model.

Although the current study has found partial answers to some of the questions that were posed, some new issues have surfaced through this analysis and they require further investigation. These matters are now canvassed briefly.

RECOMMENDATIONS

The results that have been found in this study warrant replication using other data sets obtained from the administration of attitude survey instruments. The conclusions of the present study are that item parameters are influenced by the inclusion of misfitting cases, but that the items retained through refinement using the Rasch measurement model are not. Although in this case, the refinement of the instrument was verified through the use of CFA, it seems, from the analysis of other attitude data sets not reported here, that CFA is more sensitive to response variance than is the Rasch model. That is, sets of items found to cohere under Rasch analysis do not always show such good fit under CFA. It is possible that CFA is more sensitive than Rasch modelling to this 'error variance' and that the poor model fit that emerges from CFA may be somewhat spurious. To resolve this issue, an analysis of residuals is required and through this it must be shown that item residuals are not inter-correlated in order to support the claim for unidimensionality that is implicit in the Rasch modelling. However, it is also possible that correlations among residuals can occur, not because of a failure of the unidimensionality of the item set, but because of a lack of homogeneity within the person sample. Some evidence of this was reported in the study of the CEQ (Curtis, 1999). Therefore, a more complex form of analysis that examines both the hypothesised structure of the instrument interacting with aspects of the person sample is required. For this, a form of multi-trait multi-method analysis is suggested.

In order to develop a better understanding of the effect on person-fit statistics of different pattern of responses simulated data studies are suggested. However, care must be taken to ensure that the synthetic data generated for these studies reflect the particular characteristics of attitude data. Attitude data appear to show greater variation in response choices for a given trait level, and this results in higher standard errors in both item parameter estimates and scaled

person scores. Attitudes can also be highly varied between individuals and in order to obtain reliable estimates of persons with more extreme levels of the trait, items that tap these extremities are required. However, for the majority of individuals, such items are far from their locations and responses to those items tend to be quite skewed. These skewed response patterns may influence the estimation of items parameters and case fit statistics, and in turn must influence the criterion values of those fit statistics. A study using carefully generated simulated data may shed light on the difficult question of person-fit criteria.

REFERENCES

Adams, R.J., & Khoo, S.T. (1997). Quest: the interactive test analysis system [Statistical analysis software]. Melbourne: Australian Council for Educational Research.

Bond, T.G., & Fox, C.M. (2001). Applying the Rasch model. Fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum and Associates.

Curtis, D.D. (1999). The 1996 Course Experience Questionnaire: A Re-Analysis. Unpublished Ed. D. dissertation, The Flinders University of South Australia, Adelaide.

Joreskog, K.S., & Sorbom, D. (1993). LiSRel for Windows (Version 8.12a) [Statistical analysis software]. Chicago: Scientific Software International.

Karabatsos, G. (2000). A critique of Rasch residual fit statistics. Journal of Applied Measurement, 1(2), 152-176.

Kline, P. (1993). Rasch scaling and other scales. The handbook of psychological testing. London: Routledge.

Li, M.N.F., & Olejnik, S. (1997). The power of Rasch person-fit statistics in detecting unusual response patterns. Applied Psychological Measurement, 21(3), 215-231.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. British Journal of Psychology, 88, 355-383.

Waugh, R.F. (1999). Approaches to studying for students in higher education: A Rasch measurement model analysis. British Journal of Educational Psychology, 69, 63-79.

Weiss, D.J., & Yoes, M.E. (1991). Item response theory. In R. K. Hambleton & J. N. Zaal (Eds.), Advances in educational and psychological testing: theory and applications (pp. 69-95). Boston: Kluwer Academic Publishers.

Wright, B.D., & Masters, G.N. (1982). Rating scale analysis. Chicago: MESA Press.

Wright, B.D., & Stone, M.H. (1979). Best test design. Chicago: MESA Press.

Wu, M.L., Adams, R.J., & Wilson, M.R. (1998). ConQuest generalised item response modelling software (Version 1.0) [Statistical analysis software]. Melbourne: Australian Council for Educational Research.

Adoption and implementation of information technology in Bali's local government: A comparison between single level path analyses using PLSPATH 3.01 and AMOS 4 and Multilevel Path Analyses using MPLUS 2.01

I Gusti Ngurah Darmawan
Flinders Institute of Public Policy and Management

Recent advances in information technology have aroused much interest among policy-makers, the business sector, the media and the academic world in developed countries. However, very little is known about the obstacles to accessing information technology (IT) and the diffusion and use of information technologies in developing countries, particularly the low-income economies. This research stems from the issues described above. A variety of factors may influence the outcome of adopting and implementing IT in local governments. These factors range from technological and institutional to personal, social and economic factors, and include not only the information on variables gathered at the employee level but also questions regarding the characteristics of each organization involved in the study. In addition, there are also some district characteristics collected through various secondary data sources. Hence the data files contain information obtained at three different levels, namely individual level, organizational level, and district level. In addition to identifying various factors that may affect each step of IT adoption processes, this paper also addresses more technical issues involved in determining units of analysis and choosing the proper analytical methods since the structure of the data reflects the hierarchical nature. Three different statistical packages were used to analyse the data. The exploratory path analyses were done using PLSPATH 3.01 both at individual level and organizational level separately. The results of these analyses were then compared to the AMOS 4 results. AMOS 4 could only be used to analyze the path model at the individual level due to the small number of cases at the organizational level. Based on these results, two level models were developed and further analysed with MPLUS 2.01.

technology adoption, Bali, single level, multilevel, path analysis, comparison

INTRODUCTION

In an era of globalisation, innovations in information technology, centered in telecommunications and informatics, have had substantial influence on communities and businesses. The availability of ever cheaper and more powerful personal computers, combined with the capability of telecommunication infrastructures has put increasing power into the hands of a greater number of people in organizations (Kraemer & Dedrick, 1997; Rischard, 1996; Willcocks, 1994).

The common belief is that the introduction of an innovation, especially in the form of a more advanced technology, is assumed to bring improvements to organizations (Downs & Mohr, 1976). From this perspective, IT is expected to benefit local governmental activities at operational, managerial and policy making levels (Muid, 1992, Rischard, 1996; Somers,

1987). Local governments essentially produce services rather than products, which make it difficult for them to increase productivity through automation of production processes. Rather, they should rely on efforts to improve local planning and decision making, which will in turn result in better scheduling, staffing, purchasing, and management of local programs (Kraemer & King, 1977). The common denominator of all planning and operational decisions is accurate, timely and relevant information. Most of the activities of a typical local government organization are information based in nature. Whereas most private sector organizations might consider information as a supportive input into their production processes, the government organization should see information and information handling activities as the core and essence of its operations (Taimiyu, 1993). The fact that information is regarded as a fundamental resource of government (Howard, 1985) makes this technology a suitable and prominent tool for performing governmental functions. Muid (1992, p.114) also noted that "Information systems provide the nervous system of government business. Improving the nervous system can help the body to function better". However, government organizations, both in developed and developing countries, have been slower than their private sector counterparts in adopting modern IT for harnessing their administrative and socio-economic data (Kraemer & Dedrick, 1997; Taimiyu, 1993).

The rapid changes in IT in the developed countries have serious managerial, financial, and human resource implications for information management in developing countries (Bowonder, Miyake & Singh, 1993). The shrinking of the planet into a 'global village' (Gore, 1991) makes it difficult for the developing countries to insulate themselves from the changes occurring due to developments in IT. It is thus evident that developing countries need to understand the pervasive nature of changes initiated by new IT and the consequences of not keeping pace with the changes occurring in the developed world. Therefore, the formulation and implementation of comprehensive information technology strategies have become critical for the developing countries. This is accentuated by the fact that the resources to support the operation of IT are usually scarce (Kahen, 1995; Lu & Farell, 1990; Shahabudin, 1990).

Previous research suggests that successful implementation of IT is influenced by two major factors: the technology, and the local conditions of the users. Since developing countries have distinctive environmental conditions that differ from those of developed countries and from each other, the formulation of effective strategies for IT development in a particular developing country requires a good understanding of its special environment. As they join the global information infrastructure, each country will need to find effective ways of maximizing the benefits and minimizing the risks from IT (Mansell & Wehn, 1998). Simply borrowing or transplanting from developed countries IT development experiences without a precise understanding of environmental differences is insufficient (Lu & Farrell, 1990). IT, along with many other technological innovations, will not have much operational value unless these are developed in ways relevant to the needs and conditions of the potential users, which are specific in time and place.

The importance of understanding the local conditions that might facilitate or inhibit the adoption and implementation of IT in developing countries is also stressed by the United Nation Commission on Science and Technology for Development (UNCSTD) Working Group on IT and Development, which was set up in 1995 (Mansell & Wehn, 1998).

The present research stems from the issues described above. On the one hand, there has been rapid development of IT technology and increasing needs for the technology. On the other hand, there are major macro and micro environmental differences between developed countries and developing countries and between western culture and eastern culture that need to be addressed. This study investigates the IT adoption and utilization in local

government of Bali from an innovation adoption-diffusion perspective. A four phase conceptual model of the process of adoption and implementation of innovations is used in this study namely: initiation phase, adoption phase, implementation phase, and evaluation phase (Hage & Aiken, 1970; Huff & Munro, 1985; Kumar, 1990; Panizzolo, 1998; Rogers, 1962; Rogers & Shoemaker, 1971; Zaltman, Ducan, & Holbeck, 1973).

RESEARCH MODEL

The conceptual model adopted for this study is displayed in Figure 1. A variety of factors may influence the outcome of adopting and implementing IT in local governments. These factors range from technological and institutional to personal, social and economic factors. Past research in innovation highlights the importance of human factors, organizational factors, technological factors, and the environmental factors for successful adoption and implementation of an innovation (Tornatzky & Klein, 1982; Zaltman, Duncan & Holbeck, 1973). These factors include not only the information on variables gathered at the employee level but also on questions regarding the characteristics of each organization involved in the study. In addition, there are also some district characteristics collected through various secondary data sources. Hence the data files contain information obtained at three different levels, namely individual level, organizational level, and district level.

Synthesizing various stages of innovation adoption process proposed by previous authors (Hage & Aiken, 1970; Huff & Munro, 1985; Panizzolo, 1998; Rogers, 1962; Rogers & Shoemaker, 1971; Zaltman et.al., 1973), a four phase conceptual model of innovation adoption and implementation process is formulated for this study. These four-phase innovation adoption processes consist of initiation phase, adoption phase, implementation phase, and evaluation phase. The four factors (environment, organizational, technological, and human factors) are believed to affect each phase of innovation adoption process directly or through the previous phase as the mediator.

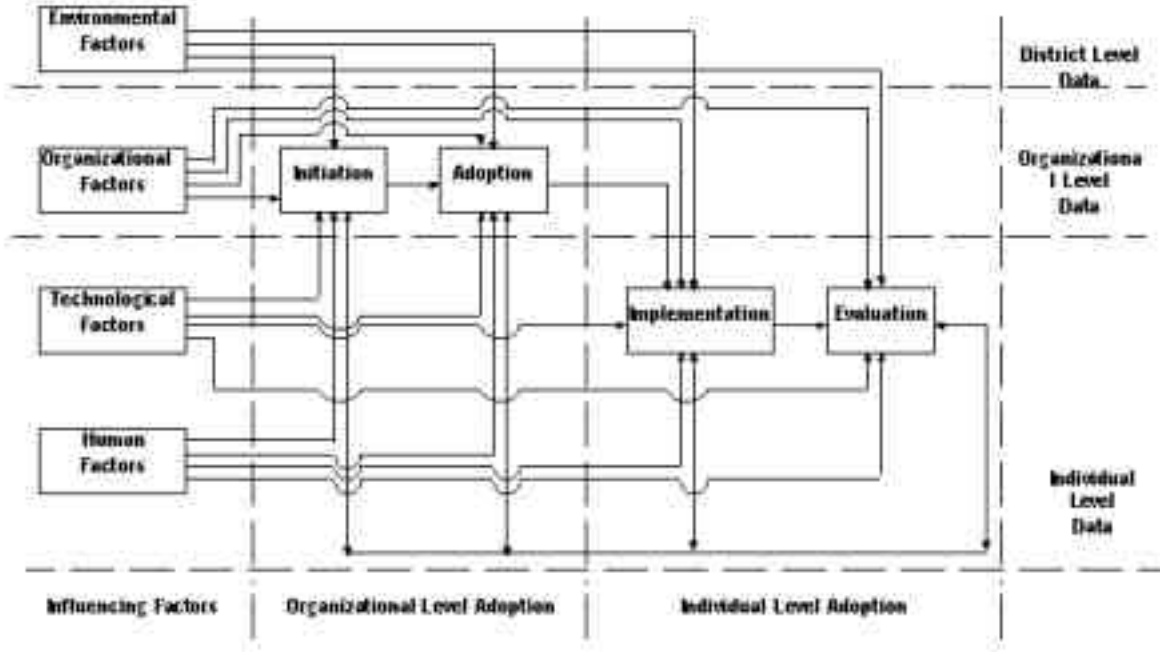


Figure 1. A Conceptual Model of IT Adoption and Utilization

This four-phase innovation adoption process can also be observed at several levels. Possible adopters of the technology are organizations, organizational units, organizational sub units, and individuals. In this study, two levels of adoption are considered. The first level of adoption, the organizational level adoption, starts when an organization begins to realize the

need for strategic change and decides to incorporate IT. It ends with acquisition of the technology. The second level of adoption, the individual level adoption, commences with the acquisition of the technology, and finishes when the technology is utilized.

However, given the constraints on time and other resources, this research study does not seek to examine all possible links. It also does not try to evaluate all aspects that are affected by the adoption and implementation of this technology. As noted earlier, there are impacts of IT on collectivities, such as the work group, the department, the organization, or even the society, and impacts on individuals. This study focuses only on the impacts of IT utilization on individuals who work in governmental agencies in term of efficiency, effectiveness, and information benefits.

In addition, there are two groups of employees involved in this study. The first group is a group of employees who were involved in the initiation stage. They took part in identifying the pressure emerging from various internal and external sources, and in recognizing the sense of urgency for a strategic change, which, in turn, led to a decision made by an organization or organizational unit that decided to incorporate IT. The second group, on the other hand, found the technology was adopted without questioning the reasons behind the acquisition of the technology. It is believed that the level of involvement in the initiation stage had significant effects on the successful rate of IT adoption and implementation processes. Therefore, two separate models were constructed to accommodate the absence of the initiation phase in the second group. For the purpose of this study the first group is called 'initiators', and the second group is called 'non-initiators', since the second was not involved in the initiation stage.

Two research models (derived from the conceptual model) proposed to be examined are shown in Figures 2 and 3 for the initiators and non-initiators respectively. The research model proposes that a number of variables in the organizations' external and internal environments, and attributes of the innovation influence the formation of pressures at the initiation phase. These variables also affect the adoption phase, implementation phase, and evaluation phase both directly or through the previous phase as a mediator.

The implementation success is a multi-dimensional concept, composed of objective and perceptual dimensions. The two dimensions of utilization and satisfaction constitute this phase and they are considered to be two major factors that influence the success of an IT implementation (Kim, Suh, & Lee, 1998). However, technology utilization and user attitude toward technology lead to performance impact. (Woodroof & Kasper, 1998; Goodhue & Thompson, 1995). Because the impacts of IT on organizations are so pervasive in post-industrial society, it is useful to define the domain of IT impacts that is the focus of this study. This study focuses on the impacts of IT utilization and user satisfaction on user performance in term of efficiency, effectiveness, and appropriateness (Kahen, 1995; Sharp, 1996; Sharp, 1998).

SOME PROBLEMS OF ANALYSIS

The research models that are to be examined in this study involve a wide variety of problematic situations. They range from various factors that may effect each step of IT adoption processes to more technical issues involved in determining units of analysis and choosing the proper analytical methods since the structure of the available data reflects the hierarchical nature of the models.

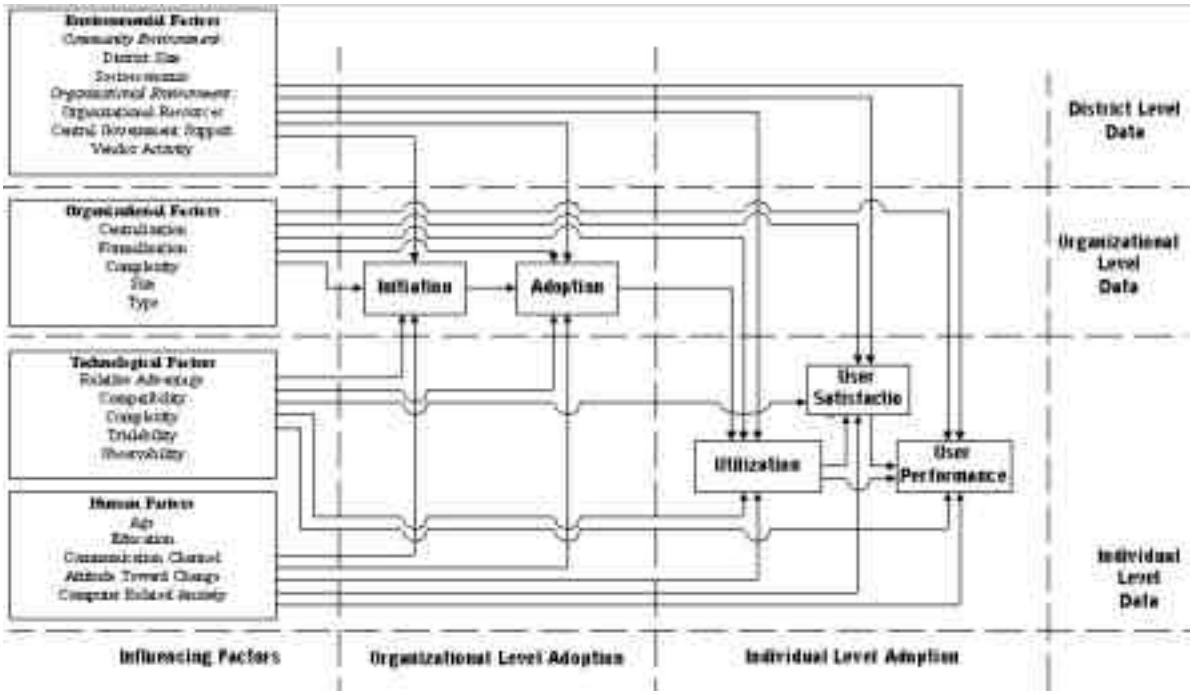


Figure 2. Research Model for Initiators

The Notion of Causality

There is one particular type of inference in this study that is necessary to make from the correlation data, namely, causal inference. The notion of causality applies whenever the occurrence of one event is reason enough to expect the production of another. Vogt (1993) gives a more precise statement to guide and restrict the application of the causality principle in theory construction and the design of research.

> To attribute cause, for X to cause Y, three conditions are necessary (but not sufficient): (1) X must precede Y; (2) X and Y must covary; (3) no rival explanations account for the covariance between X and Y. (Vogt, 1993, p.31)

The first condition appears to be met in longitudinal studies, where a sample of cases is followed over a specified period of time. However, time and financial limitations in conducting overseas field research, such as this study, do not allow such a design. With such limitations, a cross-sectional study combined with a retrospective approach to measure respondents' attitudes and perceptions before the adoption was conducted.

With regard to the second condition, an examination of correlation coefficients might indicate whether or not variation in the presumed cause is associated with the variation in the effect. With respect to the third condition, Tuijnman and Keeves (1994) have emphasized the need to specify the model under examination as precisely as possible. In addition, it is also necessary to take into consideration the principle that the causal inference begins with the assumption that any prior event might be a cause of any later event. It is, then, necessary to proceed to eliminate relations that are impossible or implausible in particular circumstances. This eliminating approach is dictated by the premise that probabilistic relations pervade both the physical and social universes and may exist even when humans are unaware of them.

Significance Testing in Social Research

Tuijnman and Keeves (1994) have emphasized the widespread but inappropriate reliance of researchers and computer programs on significance tests that assume a simple random

sample when most of the studies in social research do not follow such design. This statement is also supported by Brick at al. (1997, p.141). They say that "when data are collected as part of a complex sample survey, there is often no easy way to produce approximately unbiased and design-consistent estimates of variance analytically."

Some procedures attempt to take into consideration such sample characteristics. For example, design effects are employed to adjust for complex cluster sample design while multilevel techniques seek to take into account the nested structure of samples.

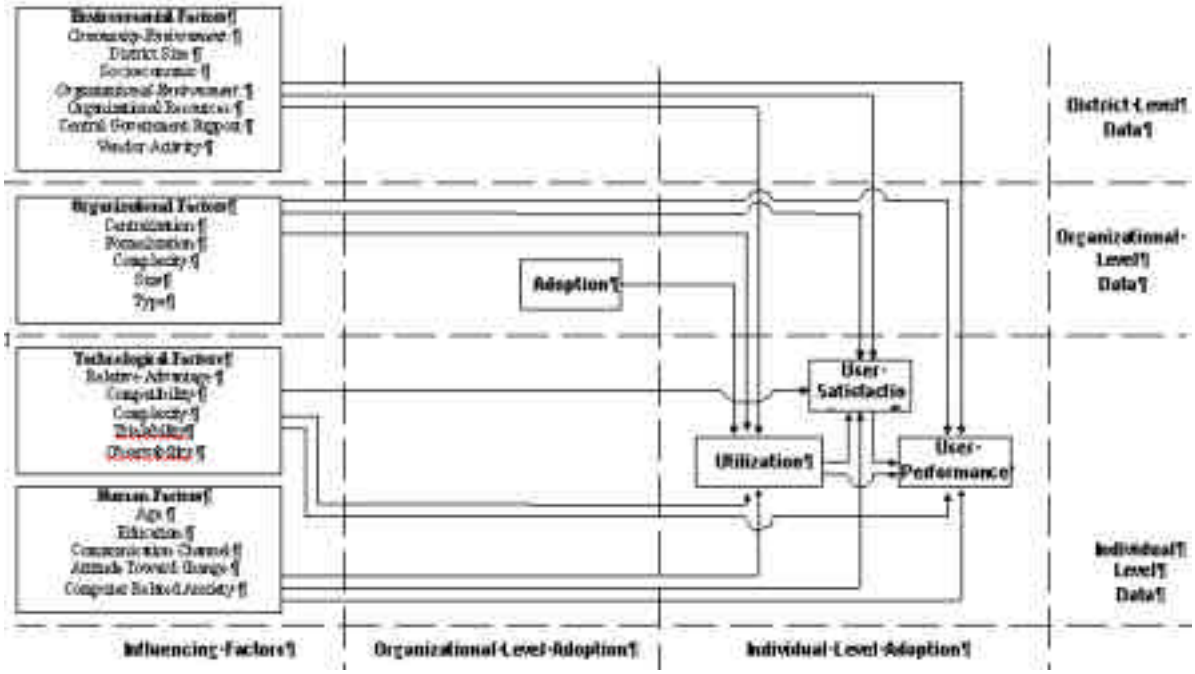


Figure 3. Research Model for Non-Initiators

Level of Analysis

It has already been pointed out that the data collected in this study included not only the information on variables gathered at the employee level but also the information on questions regarding the characteristics of each organization involved in the study. In addition, there are also some district characteristics collected from various secondary data sources. Hence, the data files contain information obtained at three different levels, namely individual level, organizational level, and district level. Models of IT adoption might be developed from theory that incorporates certain organization and district variables that might influence each phase of IT adoption processes. The examination of such models is undoubtedly of particular interest, yet severe problems arise from the inclusion of data that were obtained at different levels into one model.

Two methods that are commonly employed when data are combined from two or more levels into single-level analysis are: (a) the aggregation of data collected at the lower level (e.g. individual) to the higher level (e.g. organization); or (b) the disaggregation of higher level data to the lower level, for example by assigning organization-level data to each individual employee. Both techniques, aggregation and disaggregation, quite typically introduce bias, meaning an over- or under- estimation of the magnitude of effects associated with variables that are aggregated or disaggregated.

Aggregation bias

By aggregating individual data to the organization level, the predictive power, or fit of the regression line, commonly increases (Sellin, 1990, pp. 267-268). Furthermore, it is possible that certain variables show negative effects at one level of analysis whereas positive effects are found at another level of analysis (Sellin, 1990a, p.257). This problem is the same for all types of multivariate analysis that are confined to a single level of analysis such as ordinary least square (OLS), partial least square path analysis (PLS) or linear structural equation modeling.

Disaggregation effects

The distorting effects that occur through the disaggregation of group level data (organization and district variables) to the individual level is referred to as a disaggregation effect. In this case, the same value for a group level variable is assigned to members of the same group at the individual level. As a consequence, the assumption of the independence of observation ceases to apply. Due to the complexity of the issue, it is not possible to discuss the problem of disaggregation bias in greater detail here. Nevertheless, it is readily recognized that inappropriate error terms are estimated with data that have been disaggregated to a lower level of analysis.

DATA USED FOR THE STUDY

The data in this paper came from the study focusing on the adoption and implementation of information technology in the local government agencies of Bali-Indonesia. Bali, as a province, consists of eight regencies and one municipality. The total number of agencies that participated in this study was 153 agencies across all regions of Bali. These 153 agencies employed a total of 10,034 employees. Of these, 1,427, or approximately 14 per cent, used information technology in their daily duties. They may be considered end-users. Of these end-users, 975 employees participated in this study. The goal of this study is to examine various potential factors that might affect information technology adoption and implementation processes in the context of Bali's local government agencies. The variables used in this study are presented in Table 1.

DATA ANALYSIS

The data were analysed using three different statistical packages. The exploratory path analyses were done using PLSPATH 3.01 (Sellin, 1989). The results of these analyses were then compared to the AMOS results (Arbuckle & Wothke, 1995). Based on these results, two-level models were developed and further analysed with MPLUS 2.01 (Muthen & Muthen, 1998). The uses of these three programs are discussed in the following sections.

Partial Least Square Analyses using PLSPATH 3.0

The PLS computer program used in this study was PLSPATH version 3.01 developed by Sellin (1989) based on the Partial Least Square procedure introduced by Wold (1982) as a method for estimation of path models with latent constructs measured by multiple indicators. Sellin (1995, p.266) has claimed that PLS is "a flexible and extremely powerful technique for the examination of path models with latent constructs measured by multiple indicators." According to Sellin (1995), PLS is flexible and powerful for several reasons: it can handle big and small data files, it is technically simple, it takes little time running the program on a computer, and it does not require rigorous distributional assumptions.

Table 1. Variables in the Study

Latent Variables	Description	Manifest Variables	Description
RSIZE	District size	RSIZE	District size
SOCEC	Socio-economic level	SOCEC	Socio-economic level
SUFAC	Supporting facilities	SUFAC	Supporting facilities
OSIZE	Organizational size	OSIZE	Organizational size
TYPE	Organizational type	TYPE	Organizational type
OCOMPLEX	Organizational complexity	OCOMPLEX	Organizational complexity
CENTRAL	Level of centralization	CEN1	Centralization item 1
		CEN2	Centralization item 2
		CEN3	Centralization item 3
		CEN4	Centralization item 4
GENDER	Gender of Employee	GENDER	1=Male, 2= Female
AGE	Age of employee	AGE	Age in years
EDUC	Level of Education	EDUC	Level of Education
COMCH	Communication channel	COMCH	Communication channel
BELIEF	Belief Consistency	RDST	Reduce staff
		RDCO	Reduce cost
		CMJO	Create more jobs
		ISPL	Isolate people
		COLI	Control our live
OBSERV	Observability	SMVE	Small version
		MODU	Mudolarity
		TRYO	Try out
		SEPR	See presentation
COMPLEX	Complexity	DIUN	Difficulty in understanding
		DIUS	Difficulty in using
COMPA	Compatibility	WSCO	Work aspects compatibility
		ALCO	All aspects compatibility
		PRCO	Previous experience compatibility
RELAD	Relative advantage	INPR	Increase productivity
		INEF	Increase efficiency
		INAV	Increase availability
		INTI	Increase timeliness
		EACC	Essay to access
		SOBT	Speed to obtain
ANX	Computer related anxiety	ANX1	Anxiety item 1
		ANX2	Anxiety item 2
		ANX3	Anxiety item 3
		ANX4	Anxiety item 4
		ANX5	Anxiety item 5
ATTD	Attitude toward change	ORIG	Originator
		RULE	Rule oriented
		CONS	Conservatism
		SYST	System oriented
		FITT	Fitter in
INITIA	Initiation	NEED	Need pull
		TECH	Technology push
		SOCP	Social pressures
ADOPT	Adoption	ADOPT	Adoption
PASSIVE	Passive usage	INPA	Intensity of passive usage
		FRIN	Frequency of indirect usage
		FRPA	Frequency of passive usage
ACTIVE	Active usage	INAC	Intensity of active usage
		FRDI	Frequency of direct usage
		APPU	Application used
		TSKU	Tasks supported
SATISFAC	Satisfaction	ITSA	IT satisfaction
		SYSS	System satisfaction
		VENS	Vendor satisfaction
		TRAS	Training satisfaction
PERFORM	Performance	EFFI	Efficiency
		EFFE	Effectiveness
		APPR	Appropriateness

The use of latent variables is of great convenience because several manifest or observed variables can be grouped together to form a latent variable so that the number of paths in the

model can be reduced to a smaller number. In addition, Sellin (1995, p.256) suggests that the use of PLS is to be recommended in a condition where "the research situation at hand demands the investigation of complex models in an exploratory rather that a confirmatory fashion.The relationships among variables in the PLS path models are developed by using two sets of linear equations, referred to as the inner model and the outer model (Sellin, 1995). Sellin (1995, p. 257) defines the inner model as "the relationships between unobserved or latent variables (LVs)", and the outer model as "the relationships between LVs and their associated observed or manifest variables (MVs)." The relationships between latent variables and their associated manifest variables (outer model) are of two types. The first type is where the constructs (latent variables) are viewed as causes of measures, meaning that variations in a construct leads to variation in its measures (Edwards and Bagozzi, 2000). Such measures are termed reflective, because the represent reflections, or manifestations, of a construct. The second type is where the measures are viewed as causes of constructs (Edwards and Bagozzi, 2000). Such measures are termed formative, meaning the construct is formed or derived from its measures. Furthermore, Edwards and Bagozzi break the two types down into a more detail classifications as shown in Table 2. In this case, the two direct models are employed.

Table 2. Classification of Models Relating Constructs to Measure

Type of measure	**Reference of Measure**		
	Cause of Construct	**Attributes of Construct**	**Effect of Construct**
Reflective	Spurious model	Direct reflective model	Indirect reflective model
Formative	Indirect formative model	Direct formative model	Unanalyzed Model

Adapted from Edwards, J.R., and Bagozzi, R. P. (2000), On the nature and Direction of Relationships Between Constructs and Measures, *Psychological Methods*, v5, n2, p. 167 (pp. 155-174)

The relationship between latent variables (inner model) in the path model signify the causal relationships between the latent variables, indicate by a unidirectional arrow from the determining variable to the variable dependent on it. Keeves (1988) states that the determining variables, which do not depend on any other variables are referred to as exogenous variables, simply indicated by the fact that there is no unidirectional arrow pointing towards the variable. The variables, which are dependent on any other variables, indicated by one or more unidirectional arrows pointing towards the variables, are referred to as endogenous variables. In addition, the variable, which is dependent only on other variables is referred to as the outcome latent variable.

Structural Equation Modeling Using AMOS

Structural equation modeling (SEM) examines a series of dependence relationships simultaneously. It is particularly useful when one dependent variable become an independent variable in subsequent dependence relationships. This set of relationships, each with dependent and independent variables, is the basis of structural equation modeling. (Hair et al., 1995).

The best known structural equation modeling technique is LISREL, but this study used AMOS (Analytical of Moment Structure). The default method of computing parameter estimates in AMOS is maximum likelihood. According to Arbuckle and Wothke (1995), AMOS provides various methods for estimating structured equation models. These methods include maximum likelihood, unweighted least squares, generalized least squares, Browne's asymptotically distribution free criterion, and scale free least squares. In addition, when confronted with missing data, AMOS performs a full information maximum likelihood imputation instead of relying on the mean imputation, listwise deletion or pairwise deletion.

There are two main assumptions in structural equation modelling. First, observations must be independent. Second, the observed variables must meet certain distributional requirements. It will suffice if the observed variables have a multivariate normal distribution (Arbuckle & Wothke, 1995).

Hair et al. (1995, p.15) added, that SEM is characterized by two basic components: (1) the structural model and (2) the measurement model. The structural model is the "path" model, which relates independent to dependent variables. In such a situation, theory, prior experience, or other guidelines allow the researcher to distinguish which independent variables predict each dependent variable. The measurement model allows the researcher to use several variables (indicators) for a single independent or dependent variable. In SEM there are two kinds of variable: (1) observed and (2) latent (unobserved variable).

AMOS 4 provides two ways of controlling the program: AMOS Basic and AMOS Graphics. AMOS Basic is the macro workhorse for simplifying many specialized modelling tasks. Other object oriented programming language such as Visual Basic or C++ can also be used to control AMOS Graphics. AMOS Graphics has a graphical interface for model specification.

Critical ratios (CRs) and modification indices (MIs) are two criteria used to improve the models. The critical ratio is an observation on a random variable that has an approximately standard normal distribution. The critical ratio is obtained by dividing the estimate by its standard error. Thus, using the significance level of 0.05, any critical ratio that exceed 1.96 in magnitude would be called significant (Arbuckle & Wothke, 1995). Modification indices, on the other hand, suggest ways of improving a model by increasing the number of parameters, so that the chi-square statistic decreases faster than the degrees of freedom. However, Arbuckle and Wothke (1995, p.153) argued that in trying to improve upon a model, "a modification must only be considered if it is makes theoretical or common sense", modification indices alone should not be used exclusively as a guide.

Multilevel Path Modeling Using MPLUS 2.01

According to Keeves and Cheung (1990). five separate issues would seem to arise in single level analyses carried out using PLSPATH and AMOS. They are:

1. problems of aggregation of data from a lower to a higher level;
2. problems of disaggregation of data from a higher to a lower level;
3. specification errors which arise when a variable measured at the lower level is permitted to account for variance that is more properly associated with the higher level such as the climatic conditions in the classroom;
4. specification errors which arise when a variable measured at the higher level is not permitted to account for variance with which it is associated at the lower level, such as the variability of the regression slopes between groups; and
5. problems associated with the estimation of errors which arise with measures obtained under conditions involving two or more levels of sampling and measurement.

Given the hierarchical structure of the data, multilevel analyses are most appropriate in order to obtain satisfactory results. In this study, two-level models are analyzed using MPLUS 2.01. Initially, a level-1 model was defined based on the combined results of PLSPATH and AMOS. The models are then trimmed based on the critical ratios and modification indices in a similar way to AMOS trimming procedures. Once the level-1 models are finalized, level-2

models are defined based on PLSPATH organizational level results. The level 2 models, then, are trimmed based on critical ratios. Latent variables that estimated to have significant correlations are tried for inclusion in the model. These procedures are repeated until final results are obtained.

RESULTS AND DISCUSSION

The single level analyses results from PLSPATH and AMOS at the individual level are then compared with the 'within' model results from MPLUS. The path coefficients along with R-squared values for exogenous variables are presented in Table 5. Standard error for each MPLUS estimates is also presented in this table. For organizational level analyses, single level analyses results from PLSPATH are compared with 'between' model results from MPLUS. An attempt to do a single level analysis at the organizational level using AMOS was unsuccessful due to small sample size. The inner models comparison for organizational level analyses are presented in Table 6. In this table, MPLUS estimates are also accompanied by the standard errors of the estimates.

At the individual level, the first two analyses (PLSPATH and AMOS) are done under the assumption that each unit is independent to each other. This type of analysis is commonly referred as 'between students overall' (Keeves & Sellin, 1988). In this type of analysis, the data from different groups are pooled and a single analysis is carried out between all employees in the total sample. While the MPLUS analysis, using a two level model analysis, is more likely to employ 'between students within groups' type of analysis. In this analysis, the measures for each student are subtracted from the group mean and thus the deviation values from the group mean are employed. Moreover, the data for all groups are pooled for a combined analysis (Keeves & Sellin, 1988). In general, the 'between students overall' analysis provides more variance to be explained. This influences the R-squared values that are obtained. The variance comparison among 'between student overall', 'pooled between students within groups', and 'between groups' are presented in Table 3.

Table 3. Partitioning of Variance

Type of Analysis	Variance Component	Sum of Squares
between students overall	Due to regression	t_1
	Residual	T_1
	Total	$t_1+T_1 = t_2 + T_2 + t_3 + T_3$
between groups	Due to regression	t_2
	Residual	T_2
	Total	$t_2 + T_2$
pooled between students within groups	Due to regression	t_3
	Residual	T_3
	Total	$t_3 + T_3$

Adapted from Keeves, J. P., and Sellin, N (1988), Multilevel Analysis, in Keeves, J.P, (ed.), *Educational Research, Methodology, and Measurement: An International Handbook*, Pergamon Press, Great Britain. p. 694.

Furthermore, the significant testing for the 'between student overall' ignores the fact that the samples come from a cluster sample design. The path coefficients obtained from these analyses are overestimates. As a result, more significant path coefficients are found.

For the between group analyses, both the PLSPATH and the MPLUS analyses are employed using the between groups type of analysis. In these analyses, the individual level data are aggregated to the group level, and the group level data are then added to the model. Overall, the portions of variance explained between the two methods are somewhat similar. However, MPLUS results seem to be simpler and more meaningful.

In this study, only two levels of analyses are conducted, the individual level and the organizational level analyses. The data from the third level, district level data, are disaggregated to the organizational level data. The district level analyses are not performed due to the small number of level three units (9) and the limitation of multilevel analysis available in MPLUS which only allows two level model to be estimated. Consequently, there is still some unavoidable disaggregation bias in the results.

If it is assumed that MPLUS results are valid, 15 (54%) out of 28 relationships for initiators and 15 (44%) out of 34 relationships for non initiators are consistently estimated using PLS. Meanwhile, AMOS estimates consistently 18 (64%) out of 28 relationships for initiators and 21 (52%) out of 34 relationships at the individual level. At the higher level, 11 (37%) out of 30 relationships for initiators and 11 (33%) out of 33 relationships are consistently estimated using PLS. In this study the relationships that are regarded to be consistent are those within ± one standard error (SE) as estimated by MPLUS. In Table 4, it can be seen the percentage of path coefficients that are consistently estimated within ± 1 SE , within ± 2 SE, outside ± 2 SE, those that are not estimated by PLSPATH or AMOS but are estimated in MPLUS, and those that are estimated by PLSPATH or AMOS but are not estimated by MPLUS. In general, AMOS provides closer results to MPLUS. This may be due to the fact that AMOS and MPLUS use a maximum likelihood approach, while PLSPATH uses a least square approach with jackknife procedure to estimate the significance of each path coefficient.

Table 4. Number of Path Coefficients Estimated

Type of Relationships	**Initiators**				**Non Initiators**			
	PLSPATH		**AMOS**		**PLSPATH**		**AMOS**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Individual Level								
Estimated with coefficient within ± 1 SE	15	54	18	64	15	44	21	62
Estimated with coefficient within ± 2 SE	19	68	23	82	25	74	28	82
Estimated with coefficient outside ± 2 SE	5	18	3	11	4	12	3	9
Not estimated but exist in MPLUS	4	14	2	7	5	14	3	9
Estimated but not exist in MPLUS	23		17		13		8	
Organizational Level								
Estimated with coefficient within ± 1 SE	11	37			11	33		
Estimated with coefficient within ± 2 SE	19	63			23	70		
Estimated with coefficient outside ± 2 SE	8	27			5	15		
Not estimated but exist in MPLUS	3	10			5	15		
Estimated but not exist in MPLUS	22				13			

In the next section, the inner model results of MPLUS analyses are discussed in detail for both individual level and organizational level analyses.

Individual Level

For initiators, it would seem that the highest level of education completed is influenced positively by age. Older employees have higher qualifications. On the other hand, for non-initiators, none of the variables in the model seem to influence their level of education. Education, in turn, only influences positively communication channels used by the initiators and has no effect on non-initiators' communication channels. This result suggests that for initiators, more educated employees have more communication channels.

Table 5. Individual Level Factors: PLSPATH, AMOS and MPLUS Direct Effects Comparison

Variables	Initiators						Non-Initiators					
	PLSPATH		AMOS		MPLUS		PLSPATH		AMOS		MPLUS	
	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2
Education		0.01		0.01		0.02						
Age	0.12*		0.12*		0.13 (0.05)							
Comm. Channel				0.06		0.01		0.01		0.03		
Age			-0.21+				-0.10+		-0.10+			
Education	-		0.13*		0.12 (0.05)				0.13+			
Belief		0.05		0.06		0.02		0.01				0.02
Gender	-0.17*		-0.17*		-0.14 (0.06)		-0.09*		-		-0.14 (0.06)	
Education	-0.14+		-0.18+									
ComCh	0.11+											
Observability		0.24		0.41		0.45		0.23		0.36		0.37
Gender							-0.10*		-0.11*		-0.10 (0.05)	
Belief	0.49***		0.64*		0.67 (0.07)		0.46**		0.57*		0.58 (0.06)	
Complexity		0.15		0.28		0.28		0.17		0.27		0.25
Gender	0.08+		0.11+				0.24**		0.26**		0.13 (0.05)	
ComCh									0.16+			
Belief	0.28***		0.53*		0.53 (0.06)		0.12+					
Observability	0.17+						0.31***		0.47*		0.51 (0.06)	
Compatibility		0.34		0.40		0.37		0.28		0.34		0.39
Age												
Education	-0.09+											
Belief	0.27*		0.40**		0.29 (0.08)		0.14**		-		0.23 (0.07)	
Observability	0.38*		0.30*		0.37 (0.07)		0.40*		0.52**		0.39 (0.07)	
Complexity							0.12*		0.12*		0.12 (0.06)	
Rel. Advantage		0.30		0.34		0.39		0.30		0.36		0.33
Gender	0.15+		0.15+				0.37***		0.38***		0.20 (0.05)	
Age							0.09+		0.10+			
Belief	0.12+											
Observability	0.26*		0.32*		0.30 (0.06)		0.35*		0.39*		0.35 (0.07)	
Compatibility	0.28**		0.34*		0.40 (0.06)		0.17**		0.15***		0.29 (0.06)	
Anxiety		0.07		0.06		0.11		0.11		0.16		0.18
Belief	0.16+						0.16*		0.21*		0.21 (0.06)	
Complexity	0.16***		0.25**		0.34 (0.06)		0.26**		0.34*		0.39 (0.06)	
Compatibility							-		-0.16*		-0.16 (0.06)	
Attitude		0.12		0.17		0.18		0.11		0.15		0.17
Age									-0.10+			
Education	0.10+											
ComCh	-		0.15*		0.17 (0.05)		0.10*		-		0.11 (0.05)	
Observability							0.27**		0.32*		0.34 (0.06)	
Complexity	0.18*		0.19*		0.19 (0.06)							
Compatibility	0.26*		0.29*		0.30 (0.06)							
Anxiety							0.14*		0.15*		0.13 (0.05)	
Initiation		0.20		0.33		0.31						
Gender	-0.23*		-0.27*		-0.20 (0.07)							
Observability	0.22+		0.28+									
Complexity	0.13***		0.19***		0.40 (0.09)							
Relative Adv	-0.23+		-0.31+									
Anxiety	-0.10+		-0.14+									
Attitude	0.18*		0.26*		0.22 (0.08)							

Table 5. continued

Variables	Initiators						Non-Initiators					
	PLSPATH		AMOS		MPLUS		PLSPATH		AMOS		MPLUS	
	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2
Adoption		0.08		0.35		0.04						
ComCh	0.17+		0.28+									
Complexity	0.12+											
Relative Adv			0.20+									
Anxiety	-0.10+											
Initiation	0.13*		0.50***		0.19 (0.07)							
Passive Use		0.11		0.15		0.05		0.19		0.10		0.06
Gender							0.09**		0.23**		0.16 (0.06)	
Age	-0.18**		0.19**		0.10 (0.05)		-0.25+		0.19+			
Education			0.11+									
Belief							0.11+					
Complexity	0.13**		0.26*		0.20 (0.06)							
Observability							0.11***		0.24*		0.26 (0.07)	
Relative Adv							-		-0.17*		-0.16 (0.07)	
Attitude			0.14									
Adoption	0.22+						0.22+					
Active Use		0.37		0.49		0.11		0.40		0.47		0.34
Age	-		-0.32*		-0.27 (0.06)		-		-0.44**		-0.31 (0.06)	
Education							0.08+					
ComCh	0.16+						0.12+					
Belief	0.09+		0.14+									
Observability							0.13*		0.20**		0.11 (0.06)	
Anxiety	-0.08+											
Initiation			-0.33+									
Adoption	0.23*		0.65** *		0.16 (0.06)		0.12**		0.34***		0.18 (0.05)	
Passive Use	0.41***		0.15*		0.14 (0.06)		0.51**		0.41*		0.43 (0.07)	
Satisfaction		0.17		0.32		0.30		0.35		0.44		0.27
Gender							-		-0.11*		-0.14 (0.05)	
Belief			-0.13+				0.11+					
Relative Adv							-0.15**		-0.15**		-0.09 (0.05)	
Anxiety	0.13*		0.14*		0.14 (0.06)							
Attitude							0.22*		0.30**		0.19 (0.06)	
Initiation	0.26**		0.50**		0.36 (0.09)							
Adoption	0.15+						0.20+		0.26+			
Pasive Use	0.20+		0.14+				0.23*		0.24*		0.17 (0.07)	
Active Use	-		0.23**		0.36 (0.06)		0.19***		0.29*		0.34 (0.07)	
Performance		0.33		0.57		0.41		0.43		0.54		0.47
Gender	-0.12*		-		-0.12 (0.04)		-0.14+		-0.10+			
Age	-0.13*		-		-0.11 (0.04)		-0.11+					
Compatibility							0.12*		0.12*		0.17 (0.05)	
Relative Adv.							-0.14*		-0.16*		-0.16 (0.05)	
Attitude							0.18*		0.19*		0.15 (0.06)	
Initiation	0.14+		0.37+									
Adoption			-0.47+									
Pasive Use	0.09+						0.14+					
Active Use			0.37+				-		0.20*		0.22 (0.06)	
Satisfaction	0.46*		0.57*		0.61 (0.65)		0.41*		0.47*		0.47 (0.06)	

Note: *within ± 1 SE **within ± 2 SE ***outside ± 2 SE +estimated but not exist in MPLUS -not estimated but exist in MPLUS

Individual behaviour is deeply rooted in the person's systems of value and beliefs. Thus, if the technology were consistent with the employees' existing values and beliefs regarding the technology, it would be more likely to have positive effects on the IT adoption processes. It is found in these analyses that for both groups, consistency of beliefs is influenced by

gender. Male employees' perceptions are more consistent with the existing beliefs regarding the technology then female employees' perceptions.

Table 6. Organisational Level Factors: MPLUS and PLSPATH Inner Model Results Comparison

	Initiators				Non Initiators			
Variables	**PLS**		**MPLUS**		**PLS**		**MPLUS**	
	Beta	**R^2**	**Beta**	**R^2**	**Beta**	**R^2**	**Beta**	**R^2**
SOCEC		0.01						
RSIZE	-0.09+							
SUFAC		0.09				0.13		0.06
SOCEC	0.30+				0.35**		0.25 (0.09)	
OSIZE		0.22		0.23		0.25		0.17
RSIZE	0.44*		0.42 (0.08)		0.34*		0.37 (0.08)	
SOCEC	0.18*		0.20 (0.08)					
TYPE					0.31**		0.15 (0.08)	
CENTRAL		0.04		0.52				0.25
SOCEC	-0.19***		-0.72 (0.10)		-		-0.50 (0.11)	
OCOMPLEX		0.67		0.67		0.66		0.66
RSIZE	0.61*		0.64 (0.06)		0.66*		0.68 (0.06)	
SOCEC					0.14*		0.14 (0.05)	
OSIZE	0.35*		0.30 (0.06)		0.25*		0.24 (0.06)	
COMCH						0.09		0.03
RSIZE					0.20*		0.17 (0.10)	
TYPE					0.13+			
GENDER					0.17+			
BELIEF		0.05		0.64				1.00
RSIZE					-		0.50 (.014)	
OCOMPLEX	0.22**		0.13 (0.08)					
EDUCATION	-		-0.83 (0.09)		-		-0.74 (0.16)	
OBSERV		0.36		0.36		0.27		0.17
SOCEC	-		-0.60 (0.14)		0.23**		0.41 (0.11)	
BELIEF	0.60+				0.45+			
COMPLEX		0.24		0.20		0.36		0.34
SOCEC					-0.25+			
GENDER	0.15+				0.44*		0.59 (0.22)	
AGE	-		0.45 (0.17)					
BELIEF	0.47+							
OBSERV					0.39+			
COMPA		0.34		0.78		0.33		0.40
SUFAC					-		0.26 (0.10)	
EDUC	-0.25***		-0.88 (0.11)					
BELIEF	0.49+							
OBSERV					0.58*		0.55 (0.14)	
RELAD		0.29		0.51		0.46		0.78
CENTRAL					0.30**		0.55 (0.16)	
OCOMPLEX					-0.18*		-0.18 (0.08)	
GENDER	0.27***		0.69 (0.13)		0.45*		0.39 (0.09)	
OBSERV	0.48*		0.55 (0.16)		0.41**		0.64 (0.13)	
ANX		0.22		0.21		0.15		
CENTRAL	0.15+							
COMCH	-0.22**		-0.46 (0.12)					
BELIEF	0.38+				0.39+			
ATTD		0.25		0.44		0.21		0.09
CENTRAL	0.25+				0.20+			
OCOMPLEX					0.15**		0.31 (0.13)	
GENDER								
EDUC	0.29+							
OBSERV					0.25+			
COMPLEX	0.21+							
COMPA	0.34**		0.66 (0.18)					
ANX					0.18+			
INITIA		0.42		1.00				
GENDER	-0.34+							
OBSERV	0.35**		0.73 (0.17)					
RELAD	-0.33***		-0.66 (0.13)					
ATTD	0.33*		0.39 (0.14)					
ADOPTION		0.29		0.41				
OCOMPLEX	0.29**		0.19 (0.08)					
COMCH	0.20***		0.71 (0.13)					
RELAD	0.23*		0.13 (0.11)					
ANX	-0.18+							
INITIA	0.23**		0.44 (0.14)					
PASIVE		0.31		1.00		0.15		0.53
GENDER					0.27**		0.51 (0.14)	
AGE					0.29**		0.51 (0.14)	
OCOMPLEX	0.21*		0.27 (0.12)					
BELIEF								
COMPLEX	0.21**		0.81 (0.30)					
ATTD	0.29**		0.49 (0.22)					
ADOPTION	0.20+							
ACTIVE		0.44		0.90		0.45		0.86
RSIZE	0.20*		0.17 (0.06)		0.39***		0.85 (0.09)	

Table 6. continued

Variables	Initiators				Non Initiators			
	PLS		MPLUS		PLS		MPLUS	
	Beta	R^2	Beta	R^2	Beta	R^2	Beta	R^2
SOCEC	0.16+							
OSIZE					-0.23***		-0.44 (0.07)	
TYPE	0.17*		0.18 (0.06)					
AGE	-0.19+				-0.31*		-0.29 (0.10)	
COMCH	0.16***		0.88 (0.26)					
BELIEF								
COMPA	0.19+							
ADOPTION	0.27+				0.32**		0.17 (0.08)	
PASIVE					0.19**		0.37 (0.13)	
SATISFAC		0.25		0.68		0.45		1.00
COMPLEX					-		-0.25 (0.17)	
ATTD					0.35***		0.76 (0.20)	
INITIA	0.35***		0.83 (0.12)					
ADOPTION	0.28+				0.19***		0.42 (0.08)	
PASIVE					0.22**		0.46 (0.13)	
ACTIVE					0.33+			
PERFORM		0.46		1.00		0.52		0.90
OCOMPLEX	0.13+							
GENDER					-0.26+			
AGE					-0.17+			
COMCH	-0.25*		-0.35 (0.11)		0.18*		0.19 (0.09)	
COMPA	0.31+							
RELAD	-0.26+							
ANX					-0.20**		-0.64 (0.24)	
ATTD					0.23***		0.65 (0.20)	
SATISFAC	0.48***		0.78 (0.13)		0.50+			

Note:
*within ± 1 SE **within ± 2 SE ***outside ± 2 SE +estimated but not exist in MPLUS-not estimated but exist in MPLUS

The level of consistency of beliefs, in turn, has direct effects on the need to be exposed to the technology before adopting it. The more consistent their perceptions of the existing beliefs the higher their need to be exposed to the technology. This is a little surprising. It is expected that the more consistent their perceptions with the existing beliefs the less they need to be exposed before actually adopting the technology. This is based on the reason that if their perceptions of the technology are consistent with the existing beliefs, they tend to take the technology for granted. However, the results of this study suggest that the relationship works in the opposite direction. This kind of relationship applies for both groups. Gender, on the other hand, has negative direct effects only for the non-initiators group. The negative coefficient can be interpreted in terms that male employees have greater need to be exposed to the technology.

Employees are likely to estimate the complexity of information technology, consciously or subconsciously comparing it with the current practice or technology used. The complexity is associated with difficulty in understanding and using the new technology. It is found in this study that gender and observability have direct effects on complexity for initiators, while only belief consistency has direct effects on complexity for non-initiators. For initiators, female employees and employees with higher needs to be exposed perceive the technology to be more complex.

Extent of past experiences with computer use is expected to contribute significantly to the adoption processes. Beliefs and observability have direct effects on compatibility for initiators; while beliefs, observability, and complexity are found to have direct effects on compatibility for non-initiators.

The technology has to offer clear benefits to the organizational members in order to be adopted. The perceived relative advantages of the technology seem to be affected by gender, observability, and compatibility for initiators, and by observability, and compatibility for non-initiators in positive ways. These coefficients suggest that female employees of the initiators' group, employees that have greater need to be exposed to the technology and employees who feel they are more compatible with the technology perceived higher relative

advantages of the technology. They believe that the technology delivers greater benefits over the previous practices.

Individual characteristics and perceptions are believed to contribute to the attitudinal dimensions. In this study, it was found that computer related anxiety is affected directly by the level of consistency of beliefs, perceived complexity, and compatibility of the technology. The higher their level of consistency of beliefs and perceived complexity the more they feel anxious about the technology. The more compatible the technology the less they feel anxious. These are found for non-initiators. For initiators, only complexity is found to have a positive effect on anxiety.

Attitude toward change, on the other hand, is affected by communication channels, complexity and compatibility for initiators, and by communication channels, observability, and anxiety for non-initiators. All relationships are positive relationships. The effects of complexity for initiators and anxiety for non-initiators are rather more surprising. It is expected that, the more complex they perceived the technology, the less they want to change it. The same logical approach also applies for the effects of anxiety. However, this study revealed the opposite results. This finding indicates that non-initiators perceive the technology as a challenge rather than a deterrent.

Initiation as a latent variable only appears in the initiators' model. This is due to the fact that only this group is involved in the initiation stage. Adoption, on the other hand, appears in both models. However, in the non-initiators' model, adoption is an exogenous variable.

Gender has negative effects on initiation. This result means that female employees feel less pressures to adopt the technology. On the other hand, complexity and attitude toward change has positive effects on initiation. The effects of attitude are as expected. However, the positive effect of complexity on initiation is surprising. Therefore a further careful analysis by taking into account cross-level interactions need to be done. The level of initiation stage, in turn, has a positive direct effect on adoption.

For non-initiators, females tend to be passive users. Observability also has positive effects on passive usage in this group. On the other hand, the less they feel the relative advantages of the technology the more they tend to be passive users. For initiators, age and complexity have positive effects on passive usage as expected. Older employees and employees that perceived the technology to be difficult tend to be passive users. For active usage, age, adoption and passive usage are found to increase the level of active usage for both groups. Observability has a positive effect on active usage only for the non-initiators' group.

For the initiators, anxiety, initiation, and active usage have positive effects on user satisfaction. The more they feel the need for the technology, combined with a higher level of anxiety and a higher level of active usage, the greater the increase in the level of user satisfaction. An interesting result in this analysis is that the more they feel anxious about the technology the more they feel satisfied with it. For non-initiators, positive effects are found from attitude, passive usage, and active usage. As it is expected, this group is actually the end-users of the technology. Thus a positive effect of active usage as well as passive usage appears as expected. A negative effect is found from relative advantage and gender. This relationship indicates female employees and employees who have higher perception of the relative advantage of the technology feel less satisfied with the technology.

By observing the impact of IT on user performance in term of efficiency, effectiveness, and appropriateness, it is found for the initiators that male employees feel a greater impact of IT on their performance. The younger employees also feel the same way. Satisfaction also

works in a similar way. A higher the level of satisfaction lead to a higher level of performance. This is true for both groups. In addition, for the non-initiators, compatibility, attitude, and active usage seem to have positive effects on performance. The higher the compatibility, the higher the attitude toward change and the higher the level of passive usage, the higher the performance impact of IT. However, relative advantage has a negative effect on performance. This can be interpreted in terms that the higher they perceived the relative advantage of the technology, the less they feel the impact of the technology on their performance.

Organizational Level

From the inner model results, it is found that the availability of supporting facilities is positively correlated with the socio-economic level of the district. The higher the socio-economic level of a district, the more supporting facilities it has.

It is also found in the initiators model that the size of the organization is positively correlated with district size and socio-economic level of the district. The effect of district size is quite strong (0.42) while the socio-economic level has a moderate effect (0.20). From the non-initiators model, it is revealed that the type of the organization and district size have positive correlations with the size of the organization (0.15 and 0.31 respectively).

It is interesting that in both models, the higher the socio-economic level of a district the lesser the degree of centralization (-0.72 for initiators and 0.50 for non-initiators). This situation might be due to the fact that more developed regencies usually have higher socio-economic levels. With a larger budget available combined with a more dynamic corporate culture, organizations in those regencies experience less centralized system.

Organization complexity is reflected by the number of employees with IT skill, the level of IT skill, and the level of IT expertise of employees in that particular organization. It seems that the greater the size of an organization the more likely it has a higher complexity in terms of IT expertise (0.30 for initiators and 0.24 for non-initiators). Socio-economic level of the district is also found to be a significant factor influencing organization complexity in the non-initiators' model (0.14). One possible reason for this might be the fact that skillful employees are mostly available and have been employed in those more developed regencies.

For non-initiators, it seems that the way they communicate the ideas on IT is influenced positively by district size. On the other hand, for initiators, it seems that none of the variables in the model influence the average level of communication in each organization.

Interesting results are revealed from the effects of socio-economic level on the level of need to be exposed. In the initiators model the correlation is negative. This result suggests that for initiators, employees in organization located in a less developed district feel a greater need to be exposed while for non-initiators the relationship is the other way around.

Employees are likely to estimate the complexity of information technology, consciously or subconsciously comparing it with the current practice or technology used. The complexity is associated with difficulty in understanding and using the technology. At the organizational level, it is found in this study that gender has direct effects on complexity for non-initiators' group. People in organizations that have more female employees perceived the technology to be more complex, whereas age is found to have positive effects on complexity in the initiators' group only. This finding can be interpreted that, on average, people in the organizations that have older higher-level staff members perceive the technology to be more difficult.

Extent of the past experiences with computer use is expected to contribute significantly to the adoption processes. Education level has direct effects on compatibility for initiators, whereas supporting facilities available and observability are found to have direct effects on compatibility for non-initiators. Interestingly, people in an organization with higher average educational level of the initiators' group member feel less compatible with the technology. One possible reason for this relationship is that the higher the educational level of the employees, the more they have been exposed to the technology. Consequently the more they experience the vast change in the technology. The later technology can be very much different from the previous ones. Another possible explanation is that the higher the educational level of the employees, the more complex the tasks that they have to carry out. The limited technology resources that are available make them feel less compatible with the technology.

The technology has to offer clear benefits to the organizational members in order to be adopted. The average perceived relative advantages of the technology in an organization seem to be affected by gender and observability for initiators, and by the level of centralization, organizational complexity, gender, and observability for non-initiators in positive ways except for the complexity. These coefficients can be interpreted in terms that organizations with more female employees in both groups have higher average perception on relative advantages of the technology. They believe that the technology delivers greater benefits over the previous practices. It is also interesting to observe that the more IT expertise available in an organization, the less the average perception of relative advantages of the technology. This result suggests that those employees are the ones that realize not only the potential of the technology but also the difficulties and limitations they face in trying to delivered the benefits of the technology.

Individual characteristics and perception dimensions are believed to contribute to the attitudinal dimensions. In this study it is found that average computer related anxiety is affected directly by the level of communication channels they have. People in the organizations that communicate the idea more feel less anxious about the technology

The average attitude towards change in an organization is affected by average perceptions of compatibility for the initiators, and by organizational complexity for the non-initiators. All relationships are positive relationships as expected.

Initiation as a latent variable only appears in the initiators' model, because only people in this group are involved in the initiation stage. Adoption, on the other hand, appears in both models. However, in the non-initiators model, adoption is an exogenous variable.

Relative advantage has negative effects on initiation, while observability, and attitude toward change have positive effects on initiation. It is surprising to observe that relative advantage has a negative effect on initiation. The opposite relationship is expected. On the other hand, observability and attitude toward change have positive effects on initiation. The effects of observability and attitude are as expected. For the adoption phase, organization complexity, communication channel, relative advantages and initiation have positive effects as expected.

For non-initiators, organizations with more female and older employees are tent to have a higher level of passive usage. For initiators, organizational complexity, IT complexity, and attitude toward change act on passive usage as expected. District size increases the level of active usage accordingly for both groups. For initiators, in addition, organizational type and communication channels have positive effects on active usage, whereas for non-initiators, organizational size has a negative effect. This may be related to the observation that for large organizations it can be expected that the average active usage in the organization is lower.

Many low level staff members have limited opportunities to use the technology actively due to the limited amount of technology available for them. However, it is also found that for the non-initiators, a high level of average passive usage and adoption in an organization lead to a higher level of average active usage and people in the organizations that have younger employees tent to have higher level of active usage.

For the initiators, initiation has positive effects on user satisfaction. The more they feel the need for the technology the higher their level of satisfaction. For non-initiators, positive effects are found from attitude, adoption, and passive usage.

It is also found in this study that for the initiators, communication channels and satisfaction are effecting user performance. A high level of satisfaction, as expected, leads to a high perception of IT impact on performance. In contrast, communication channels have negative effects on user performance. This relationship can be interpreted in terms that in the organizations where they communicate the ideas regarding IT intensively people tend to feel less impact of the technology on their performance. However, communication channels are also found to work in the opposite way for the non-initiators. This result suggests that for low level staff, the actual end users, the more they communicate, the more they feel the impact of IT on their performance. In addition, for non-initiators, anxiety is found to have negative effects on performance and attitude toward change operates in the opposite direction.

The relationships among variables at the individual level and the organizational level are presented in Figure 4 and Figure 5 for initiators and non-initiators respectively.

In order to identify facilitators and inhibitors for each phase of IT adoption processes, summary results are presented in Table 7 and Table 8 for individual level model and organizational level model respectively. In addition to the direct effects, those tables also provide significant indirect effects, those with indirect coefficients of 0.1 or higher, that influenced each phase of IT adoption processes.

Table 7. Facilitators and Inhibitors at the individual level

	Initiators		**Non-Initiators**	
Criterion	**Facilitators**	**Inhibitors**	**Facilitators**	**Inhibitors**
Initiation				
Direct effect	Complexity, Attitude	Gender		
Indirect effect	Belief, Compatibility			
Adoption				
Direct effect	Initiation			
Indirect effect				
Passive Usage				.
Direct effect	Age, Complexity		Gender, Observability	Relative advantages
Indirect effect	Belief		Belief	
Active Usage				
Direct effect	Adoption, Passive Usage	Age	Anxiety, Adoption, Passive Usage	Age
Indirect effect			Observability	
User Satisfaction				
Direct effect	Anxiety, Initiation, Active usage		Attitude, Passive usage, Active usage	Gender, Relative advantages
Indirect effect	Complexity	Age		Age
Impact on Performance				
Direct effect	Satisfaction	Age, Gender	Compatibility, Attitude, Active Usage, Satisfaction	Relative advantages
Indirect effect	Anxiety, Initiation, Active usage			

Table 8. Facilitators and Inhibitors at the Organizational level

	Initiators		**Non-Initiators**	
Criterion	**Facilitators**	**Inhibitors**	**Facilitators**	**Inhibitors**
Initiation				
Direct effect	Observability, Attitude	Relative adv.		
Indirect effect	Compatibility	Socio- eco.		
Adoption				
Direct effect	Org. Complexity, Comm. Channel, Relative Adv., Initiation			
Indirect effect	District Size, Observability, Attitude	Relative adv.		
Passive Usage				.
Direct effect	Org. Complexity, Complexity, Attitude		Gender, Age	
Indirect effect	District size, Age, Compatibility	Education		
Active Usage				
Direct effect	District size, Org, Type, Comm. Channel		District Size, Adoption, Passive usage	Age, Org. Size,
Indirect effect			Gender	
User Satisfaction				
Direct effect	Initiation		Attitude, Adoption, Passive Usage	Complexity,
Indirect effect	Obervability, Attitude, Compatibility	Socio-Ec., Gender, Education, Observability Relative Adv.	Org. Complexity, District Size, Gender, Age	
Impact on Performance				
Direct effect	Satisfactionion	Comm. Channel	Comm. Channel, Attitude,	Anxiety
Indirect effect	Initiation, Observability, Attitude, Compatibility	Gender, Relative adv., Education	Org. Complexity, District Size	

CONCLUSION

The findings of this study contribute to both the theoretical and empirical knowledge on organizational adoption and implementation of IT innovations for those governmental agencies in developing countries by adding to the case studies available the local characteristics of Bali, Indonesia. In addition, this study also provides a contribution in identifying the facilitators and inhibitors for IT adoption and implementation in local government agencies of Bali. These facilitators and inhibitors at the individual level and at the organizational level are presented in Table 7 and Table 8 respectively.

By recognizing these factors, Bali's government agencies are expected to be able to formulate better strategies in adopting and implementing IT in order to increase their service quality and productivity. This could help Bali tourism to compete with other tourist destinations in a highly competitive world-wide tourism industry. It also provides the Indonesian government with a better understanding of local conditions in Bali for formulating its IT policy.

In addition, on the methodological side, it is also confirmed that analysing hierarchical data with separate single level analyses at different levels suffer from aggregation and disaggregation effects. Consequently, multi-level analyses were needed in order to obtain satisfactory results. However, in order to analyses a complex model, a series of single level exploratory analyses are recommended. Among other software packages, PLSPATH and AMOS can be used for this purpose.

PLSPATH has major advantages when dealing with small sample size and may categorical variables. In addition, PLSPATH can handle two types of relationships between latent variable and the associated observed variables, inward mode and outward mode. The SEM packages like AMOS and LISREL are not designed to handle the inward mode. It should be noted that one of the limitations of the use of the PLSPATH program is the absence of standard errors. However, Sellin (1990) has indicated that the estimates of the standard errors using the jackknife method can be employed for examining the significance of the result. In addition, the fit indices for the indirect effects can indicate the removed paths that should be put back in the inner model.

AMOS and MPLUS, on the other hand, both provide the test of significance and goodness of fit indices to indicate the goodness of the model to fit the data. AMOS and MPLUS also have the capability to correlate measurement errors. This can decrease the RMSEA indices to make the model fit better with the data. In addition, AMOS has a major benefit in visual modelling. Using its graphical interface makes model building easier.

In addition, on the methodological side, it is also confirmed that analysing hierarchical data with separate single level analyses at different levels suffer from aggregation and disaggregation effects. Consequently, multi-level analyses were needed in order to obtain satisfactory results. However, in order to analyses a complex model, a series of single level exploratory analyses are recommended. Among other software packages, PLSPATH and AMOS can be used for this purpose.

PLSPATH has major advantages when dealing with small sample size and may categorical variables. In addition, PLSPATH can handle two types of relationships between latent variable and the associated observed variables, inward mode and outward mode. The SEM packages like AMOS and LISREL are not designed to handle the inward mode. It should be noted that one of the limitations of the use of the PLSPATH program is the absence of standard errors. However, Sellin (1990) has indicated that the estimates of the standard errors using the jackknife method can be employed for examining the significance of the result. In addition, the fit indices for the indirect effects can indicate the removed paths that should be put back in the inner model.

AMOS and MPLUS, on the other hand, both provide the test of significance and goodness of fit indices to indicate the goodness of the model to fit the data. AMOS and MPLUS also have the capability to correlate measurement errors. This can decrease the RMSEA indices to make the model fit better with the data. In addition, AMOS has a major benefit in visual modelling. Using its graphical interface makes model building easier.

References

Arbuckle, J. L. and Wothke, W. (1999). *Amos 4.0 User's Guide*, SmallWater Corporation, USA

Bowonder, B., Miyake, T., and Singh, T. M. (1993). Emerging Trends in Information Technology: Implications for Developing Countries, *International Journal of Information Management,* v13, pp. 183-204

Brick, J.M., Broene P., James, P., and Severynse, J. (1996*). A User's Guide to WesVarPC (Version 2.0).*, MD, Westat Inc., Rockville

Downs, G. W. and Mohr, L.B. (1976). Conceptual Issues in the study of Innovation, *Administrative Science Quarterly*, v21, pp.700-714

Edwards, J.R., and Bagozzi, R. P. (2000). On the nature and Direction of Relationships Between Constructs and Measures, *Psychological Methods*, v5, n2, pp. 155-174

Goodhue, D.L., and Thompson R.L. (1995). Task-Technology Fit and Individual Performance, *MIS Quarterly*, June, pp. 213-236

Gore, Al (1991). Computer, Networks, and Public Policy: Infrastructure for the Global Village, *Scientific-American*, September, pp. 150-153

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black W.C. (1995). *Multivariate Data Analysis*, Prentice Hall Englewood Ciffs, New Jersey

Hage, J. and Aiken, M. (1967). Program Change and Organizational Properties: A Comparative Analysis, *American Journal of Sociology*, v72, pp. 505-519

Howard, E.W. (1985). New Technology and Organizational Change in Local Authorities, *Computers, Environment and Urban Systems*, v10, n1, pp. 9-18

Huff, S.L., and Munro, M.C. (1985). IT Assessment and Adoption: A Field Study, *MIS Quarterly*, Dec, pp. 327-339

Kahen, G. (1995). Assessment of Information Technology for Developing Countries: Appropriateness, Local Constraints, IT Characteristics and Impacts, *International Journal of Applications in Technology*, v8, n5/6, pp. 325-332

Keeves, J. P., and Sellin, N (1988). Multilevel Analysis, in Keeves, J.P, (ed.), *Educational Research, Methodology, and Measurement: An International Handbook*, Pergamon Press, Great Britain

Kim, C., Suh, K., and Lee, J., 1998. Utilization and User Satisfaction in End-User Computing: A Task Contingent Model, *Information Resources Management Journal*, v11, n4, pp. 11-24

Kraemer, K.L., and Dedrick, J. (1997). Computing and Public organizations, *Journal of Public Administration Research and Theory*, v7, n1, pp. 89-113

Kraemer, K.L., and King, J.L. (1977). *Computers and Local Government*, Praeger Publisher , New York

Kumar, K. (1990). Post Implementation evaluation of computer-based Information Systems: current practices, *Communication of the ACM*, v33, n2, pp. 203-212

Lu, Ming T. and Farrell, C. (1990). Information Systems Development in Developing Countries: An Evaluation and Recommendation, *International Journal of Information management*, v10, pp. 288-296

Mansell, R. and Wehn, U. (1998). Knowledge Society: Information Technology for Sustainable Development, Oxford University Press, New York

Muid, C. (1994). Information Systems and new public management – a view from the center, *Public Administration*, v72, n1, pp. 113-126

Muthen, L. K. and Muthen B. O., (1998). *Mplus User's Guide: The Comprehensive Modeling Program for Applied Researchers*, Muthen & Muthen, California

Panizzolo, R. (1998). Managing Innovations in SMEs: A Multiple Case Analysis of the Adoption and Implementation of Product and Process Design Technologies, *Small Business Economics*, v11, pp. 25-42

Rischard, J. F. (1996). Connecting Developing Countries to the Information Technology Revolution, *SAIS Review*, Winter-Spring 1996, pp. 93-107

Rogers, E.M.(1962). *Diffusion of Innovations*, The Free Press, New York

Rogers, E.M.(1983). *Diffusion of Innovations*, 3rd ed., The Free Press, New York

Rogers, E.M. and Shoemaker, F.F. (1971). *Communication of Innovations*, The Free Press, New York

Sellin, N. (1989). PLSPATH version 3.01 Application Manual, Hamburg, Germany, Humberg University

Sellin, N. (1990). On Agreegation Bias, In K.C. Cheung, J.P. Keeves, N Sellin, and S.C. Tsoi, The Analysis of Multilevel Data in Education Research: Studies of Problems and Their Solutions, *International Journal of Education*, v14, pp. 257-268

Sellin, N. (1995). Partial Least Squares Modeling in Research on Education Achievement, In W. Bos and R.H. Lehmann (Eds), *Reflections on Education Achievement*, Munster Germany, Waxmann Verlag, pp. 256-267

Shahabudin, S. (1990). Information Technology in Developing Countries, *International Journal of Computer Applications in Technology*, v3, n3, pp. 194-200

Sharp, C.A. (1996). *Designing Your Own Goal Attainment Measures for Evaluation*, P.E.R.S.O.N.A.L. (Research & Consultancy) Pty. Ltd., South Australia

Sharp, C.A. (1998). Public Sector Performance Measurement, Paper presented at the *IIR Conferences*, Perth, June

Somers, R. (1987). Geographic Information Systems in Local Government: A Commentary, *Photogrammetric Engineering and Remote Sensing*, v53, n10, pp. 1379-1482

Tiamiyu, M. A. (1993). The realities of developing modern information resources management system in government organizations in developing countries with particular reference to Nigeria, *Journal of Information Science*, v19, pp. 189 – 198

Tornatzky, L.G., and Klein, K.J. (1982). Innovation Characteristics and Innovation Adoption-Implementation: A Meta Analysis of Findings, *IEEE Transactions on Engineering Management*, v29, n1, pp. 28-45

Tuijnman, A.C. and Keeves, J. P. (1994). Path Analysis and Linear Structural Relations Analysis, In T. Husen and T.N. Postlwtwaite (Eds*.), The International Encyclopedia of Education* (2nd edn), Oxford, Pergamon Press, pp. 4339-4352

Vogt, W. P. (1993). Dictionary of statistics and methodology, A nontechnical guide for the social sciences, Newbury Park, Sage Publications, California

Willcocks, Leslie (1994). Managing Information Systems in UK public administration, *Public Administration*, v72, n1, pp. 13-32

Wold, H. (1982). Soft Modeling: The Basic Design and Some Extentions, In K.G. Joreskog and H. Wold (Eds), *System Under Indirect Observation: Part III*, Amsterdam, North Holland Press, The Netherlands

Woodroof, JAB. and Kasper, GM (1998). A Conceptual Development of Process and Outcome User Satisfaction, *Information Resources Management Journal*, v11, n2, pp. 37-43

Zaltman, G., Duncan, R., and Holbek, J. (1973). *Innovation and Organizations*, John Wiley and Sons, New York

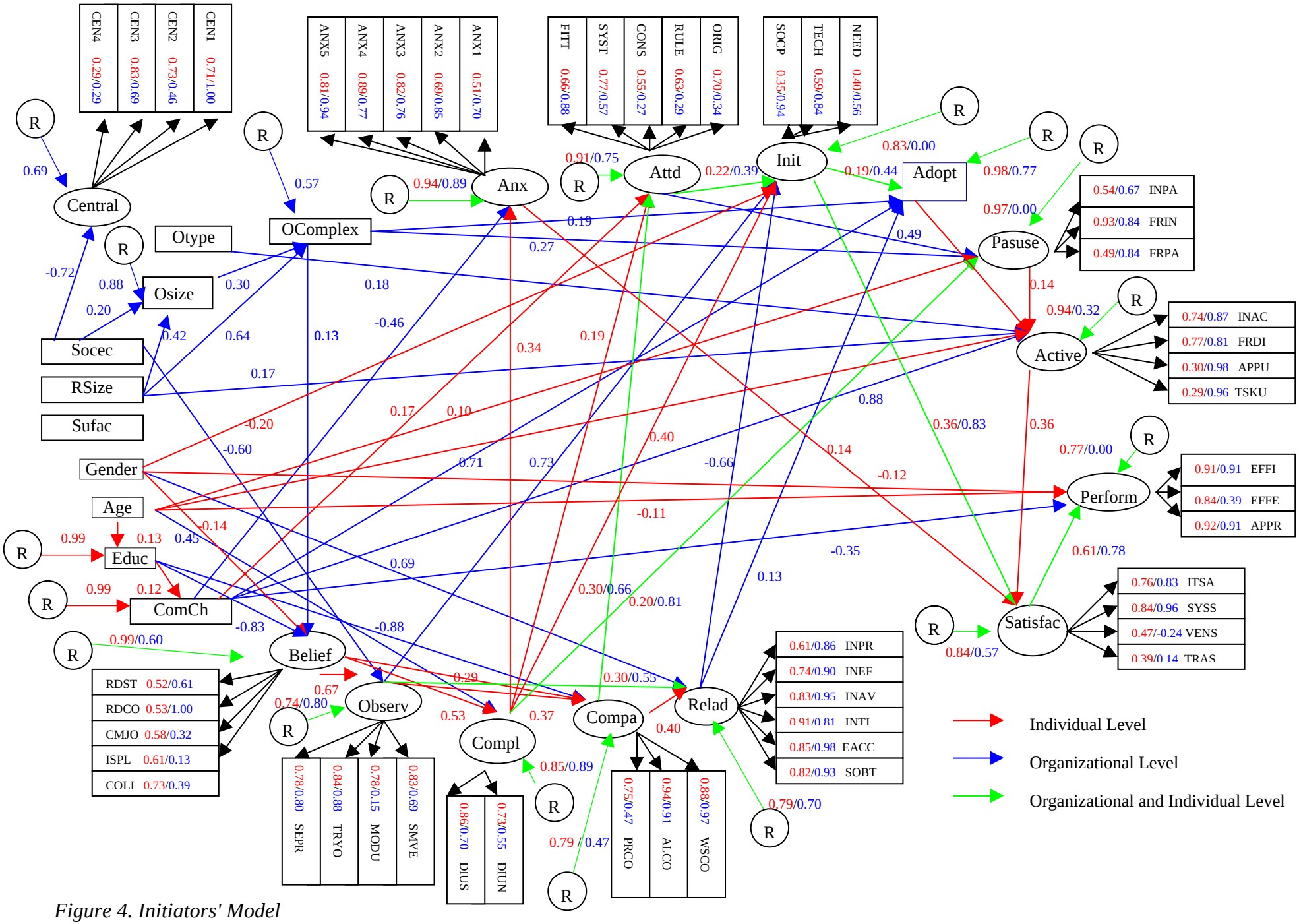


Figure 4. Initiators' Model

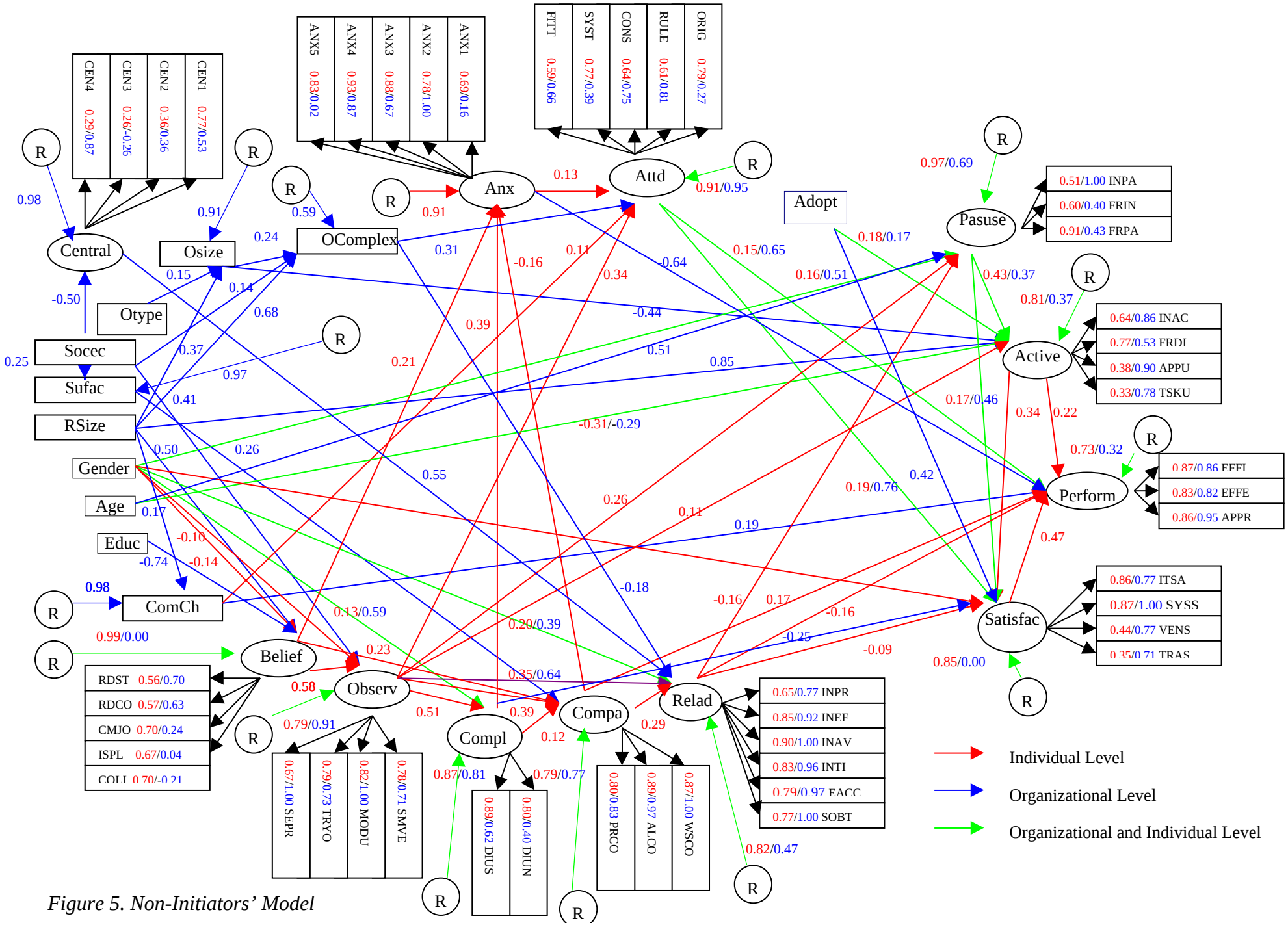


Figure 5. Non-Initiators' Model

The globalisation of the English language: Reflections on the teaching of English in China

Du Hui
Flinders University of South Australia *duhuig@hotmail.com*
China Pharmaceutical University; Nanjing, China

Using data collection from schools and government documents, the paper discusses the switch from the teaching of Russian to that of English, presents a picture of English teaching in China at present, and predicts some of the changes in the teaching of English in the future in this country. The paper focuses on the following issues: a) Why has China chosen English as the foreign language most widely taught in schools? b) What is the influence of globalisation of the English language in China? c) Why is the teaching of English important to China? The paper consists of three parts: Stages of foreign language teaching since 1949; The situation of English teaching in China; Looking at English teaching from a cultural perspective.

globalisation, English teaching, College English, international language, cultural perspective

INTRODUCTION

Globalisation is an influencing force everywhere in the world. It makes us more aware that, along with the prescriptive slogan 'think globally, act locally', there must, as so aptly put by Hogan (1996), be the additional critical observation that we also 'think locally, act globally'. China has the world's largest population and a history of over 5000 years. In recent years China has chosen English as the major foreign language taught in all colleges and universities as well as in almost all high schools. English is taught from Grade 5 in many elementary schools, and in some, even from Grade 3. In every large city there is at least one high school of foreign languages, which attaches great importance to English teaching. For those who have had the experience of learning English from their elementary schools to universities, the time and energy they have spent on English language learning is greater than that on Chinese language learning. While people in China pay attention to the English language, they also have started to become interested in aspects of western culture. Even in the mid-1980s, people sent cards to each other mostly before New Year, saying 'Happy New Year!'. But since the 1990s, people have started to send cards to each other before Christmas, saying 'Merry Christmas and Happy New Year!' English not only appears in media, but also in people's daily conversations as well as on their clothes, and provide examples of English integrating into traditional Chinese culture. In other words, Chinese culture is enriched with the introduction of English culture. Why is English as a foreign language the most widely taught in China? What is the relationship between the globalisation of the English language and China? In the age of globalisation, it seems necessary to discuss these issues from a cultural perspective.

STAGES OF FOREIGN LANGUAGE TEACHING IN CHINA

Russian teaching

Chairman Mao (1949, p.1471) once said: 'The firing of the cannons in Russia in October 1917 sent Marxism to China', which shows that Marxism in China comes from Russia. Since Marxism came to China, there has been a very special relationship between China and the

former Union of Soviet Socialist Republics (USSR). The Brotherhood relationship, from 1921 to 1959, was followed by a period of hostility, between 1960 to 1988, which moderated into the present day relationship. The relationship between the two countries has actually been between the Soviet Communist Party and the Chinese Communist Party and dates back to 1921 when the Chinese Communist Party was founded.

After World War II, the whole world was divided into two groups: that of the Capitalist Group, centered around the USA, and the Socialist Group, centered around the USSR. After the founding of the People's Republic of China, the Chinese Communist Party decided to build China into a socialist country. But the Chinese Communist Party had no such experiences. Since USSR, the centre of the Socialist Group, had already gone on its road of socialism and there had been a very close relationship between the Soviet Communist Party and the Chinese Communist Party, it was quite natural for the Chinese Communist Party to rely on the USSR. As a result, China's policy-making was greatly influenced. The Chinese Communist Party and government sent many leaders and intellectuals to the USSR to learn from the Russians. However, as a proportion of the population in China, those who had the chance to go there were only very few in number. To learn from the USSR, more and more people needed to be able to know the Russian language and communicate with Russians. It seemed that the only way to achieve such a goal was to teach Russian in schools. Thus, very soon Russian became the most widely taught foreign language in schools: not only in colleges and universities but also in high schools. Russian was promoted at the expense of English because of China's political alliance and economic construction (Hu, In Press). At that time, people took great pride in learning the Russian language. In China Pharmaceutical University there is a veteran professor of English, who entered Beijing Foreign Languages Institute in 1952 with the desire to learn Russian. However, the school arranged for her to learn English. She felt rather disappointed because Russian was then regarded as communism's language while English was considered the imperialism's language. After much persuasion by her teachers, she finally agreed to learn the English language.

Foreign language teaching: More absent than present

The relationship between the Soviet Communist Party and the Chinese Communist Party began worsening after the death of Stalin in 1953. In 1960, the USSR withdrew all of its experts from China and stopped all aid projects and contracts with China. This shocked the Chinese Communist Party greatly. The Chinese government realised that China should build socialism by itself, while criticising the USSR as a revisionist country rather than a socialist one, China regarded herself as the centre of the Socialist Group. China was then in a situation that it could neither learn from the Western world − the Capitalist Group − nor from the USSR for its being a revisionist country. In 1966, the Cultural Revolution broke out. Almost everything foreign, traditional, or so-called revisionist was strongly criticized. Consequently, foreign language teaching was almost completely stopped. To be exact, it was more absent than present from 1967 to 1976, during which period if one listened to a foreign broadcasting station, one would be suspected as a spy or special agent and would be investigated, or even arrested and put into prison.

English teaching

When the political agenda prevails over the educational, foreign language education suffers and when the political agenda converges with the educational, foreign language education gains (Hu, In Press). From 1966 to 1976, higher education in China existed in name only. China had almost completely closed its doors to the outside. It was during those 10 years that the gap in science and technology between China and other countries, such as Japan and USA, was greatly widened. When looking through the narrow opening of its door 10 years later, the Chinese government seemed to be shocked by the rapid developments of other countries as well as the impact of the globalisation of the English language. To improve its backward position and catch

up with advanced countries, something had to be done. As a result, China stopped enrolling university students from the peasants and workers by means of recommendation and resumed the College Entrance Examination in 1977 (Deng, 1977). In 1978, the Third Plenary Session of the Eleventh Central Committee of the Chinese Communist Party was held in Beijing, the motto of which was 'Practice is the only real criterion to test the truth'. As a turning point in the development of China, the meeting decided that economic construction should be the central task of the whole country and China should reform and carry out the opening-to-the-outside-world policy (Deng, 1978). As a precondition of implementing this policy, there should be people in China who know foreign languages. Deng Xiaoping (1979) required that all people who were able to learn a foreign language should learn it. In the situation that English was becoming an international language, only by knowing English would Chinese people communicate with most of the other peoples in the world and access the most advanced science and technology. Such a reality led to the policy of making the teaching of English as the major foreign language in China.

THE SITUATION OF ENGLISH TEACHING IN CHINA

English teaching in schools

English is the most widely taught foreign language in China. In municipalities directly under the central government and capital cities of every province as well as in some developed cities, English is taught from Grade 5 in elementary schools, and in some, from Grade 3. Usually, there are three lessons of English every week. *People's Education Press* and *Longman* mainly publish the English textbooks used in elementary schools. Never before have parents in China paid so much attention to their children's education. Under the persuasion or pressures of their parents, nearly every child has to choose something special to learn, such as painting, singing, swimming, playing musical instruments and so on, but English is what most children are learning. Most parents spend much money on employing tutors to teach their children English, whether their children have already started to learn English formally in school or not. These parents hope that in this way their children can be good at English or at least when the children start to learn English in school formally they will learn it easily.

Recently, it is reported that English will be taught from Grade 3 in every elementary school in China. Except in very remote and backward places, English is taught in every junior and senior high school. The Ministry of Education is in charge of compiling the English textbooks used in high schools. Both in junior high school and in senior high school, English is divided into two bands: Band 1 and Band 2. For each Band, there are specific requirements. For example, according to the English Syllabus for Junior High School (2000), students are required to meet the specific requirements for Band 1 in the following five aspects:

a) Vocabulary: to be familiar with 450 words and 100 expressions.
b) Reading: to read at the speed of 40~50 wpm (words per minute).
c) Listening: to understand the listening materials at the speed of 100 wpm.
d) Writing: after listening to it for three times, students are required at the speed of 6~8 wpm to dictate the text which they have learned and is now written.
e) Speaking: to be able to ask and answer simple questions on the text as well as make simple conversations on familiar topics in daily communication.

In junior high school, both English and Chinese are taught five periods per week respectively; while in senior high schools, both English and Chinese are taught six or seven periods per week respectively. It appears that English and Chinese are taught for the same length of time in high schools. Compared with the pupils in elementary schools, students in junior and senior high schools have to bear more pressures from the English course because English is the subject to be examined both in senior high school entrance examination and college entrance examination.

In every college or university, English is a compulsory course for every non-English major student. Such an English course is called College English in China. College English is divided into six bands: from Band 1 to Band 6, which students should complete in six semesters. That is to say, one Band each semester. The two requirements of the College English Teaching Syllabus (1991) are the basic requirement and the higher requirement. Accordingly, at the basic level, students should meet the specific requirements in the following six aspects:

a) Vocabulary: to recognize 4,000 words, of which they are to be familiar with 2,300 words.
b) Pronunciation: to read aloud with acceptable pronunciation and intonation.
c) Reading: to employ basic reading skills and read with ease texts on general topics at intermediate level. The reading speed should be no less than 50 wpm with at least 70 per cent accuracy in comprehension.
d) Listening: to follow classroom instruction in English and to understand listening materials on familiar topics presented only once at the speed of 120 wpm with the score no less than 70 per cent accuracy in comprehension.
e) Writing: With some guidance, to complete in 30 minutes a short composition of 100 - 120 words on a given topic.
f) Speaking: to carry on daily conversations in simple English and to ask and answer questions based on a given text.

Every student in the key universities must meet the basic requirement, which can be shown by whether a student has passed the College English Test Band 4 (CET 4) or not. Such a requirement was originally only for students in key universities, but other universities have followed. They also require that each of their students pass CET 4. Those who fail to pass CET 4 may not, for example, gain their bachelor degrees, and in some schools are not even awarded their diplomas. Once a student passes CET 4, there is no compulsory requirement for him/her to pass College English Test Band 6 (CET 6). However, most of those who have passed CET 4 will go on trying to pass CET 6: some succeed, some fail. In 1999, China started the CET Spoken English Test. Those whose scores are above 85 are entitled to participate in such a test. Nearly every student who meets the requirement takes part in the test.

There are four or five periods of every week. In every school year, College English Test (CET) takes places twice nationally: one is in the middle of June while the other is in late December or middle of January. At each time, the CET consists of two kinds: CET 4 and CET 6. There is an organisation called the CET Committee in charge of the testing, under which there are three centres: Beijing, Shanghai and Wuhan. The CET Committee is directly under the Ministry of Education. According to authoritative statistics (1999), students participating in such examinations range from 100,000 at the beginning in 1987 to 2,400,000 every year from 1997 to 1999.

Table 1 shows the situation in which the Chinese and English Languages are taught in elementary schools, junior high schools, senior high schools and universities, respectively. We can see that the periods of the Chinese language and those of the English language are similar. However, Chinese students have to spend much more time on learning English than Chinese because it is more difficult to learn a foreign language than to learn one's native language.

Table 1. Teaching of both Chinese and English in Chinese schools

Schools	**The Chinese Language**		**The English Language**	
	Grades	Periods/Week	Grades	Periods/Week
6-Year Elementary School	1~6	6	5~6	3
3-YearJunior High School	1~3	5	1~3	5
3-Year Senior High School	1~3	6~7	1~3	6~7
4-Year University	0	0	1~2	4~5

English teaching outside schools

In addition to English teaching in schools, there is also English teaching outside schools. The students are usually adults from all walks of life. They study English for different special purposes. In China, any one who wants to get a promotion in his/her specialty must pass a certain level of English test. Once you pass the test, you will get a certificate usually valid for three years. However, if you do not get your promotion within the 3 years, you will have to pass the English test of the same level again. Learning English for promotion is one of the most important reasons for adults to learn English. People also learn English for other purposes, like doing business with foreign countries, going abroad to tour, study or to live, and so on. Thus, some organizations have appeared to meet such needs. Well-known teachers of English are usually employed to finish the task-based teaching by these organizations, some of which are running so well that they are nationally famous, such as 'Beijing New Orientation School' in Beijing, 'Qianjin Spare-Time School for Advanced Studies' in Shanghai and 'Jingling Translation Society' in Nanjing. Therefore, competent teachers of English are always popular and never lack money.

LOOKING AT ENGLISH TEACHING FROM A CULTURAL PERSPECTIVE

Globalisation of the English language and China

Neave (2000) says that for the past quarter century, higher education has been high on the agenda of governments and central to the fortune of nations. That China has chosen English as a foreign language to be taught not only in higher education but also in secondary and elementary education is also central to the fortunes of China. Why is English taught in China? Carl Marx once said that if you could sing *The Internationale*, you would be able to find friends everywhere in the world. It is true of English. Crystal (1997) reports that 85 per cent of international organisations make official use of English, at least 85 per cent of the world's film market is in English, and about 90 per cent of published articles are written in English. In addition, Zai, Zheng and Zhang (1999) report that more than 80 per cent of all scientific papers are first published in English and over half of the world's technical and scientific periodicals are in English. Furthermore, five of the largest broadcasters (CBS, NBC, ABC, BBC and CBC) reach a potential audience of about 300 million people through English broadcasts, computer program instructions and software are often supplied only in English, 85 per cent of the international telephone conversations are conducted in English, and English is the language of medicine, electronics, space and the Olympics.

From what has been mentioned above, we might come to such a conclusion: English does not only belong to English speaking countries, it belongs to the whole world. No other language is more widespread geographically than the English language. China is carrying out an opening-to-the-outside policy so that it can better know the world and communicate with the world. However, without knowing English, it cannot know the world. In summary, China has adopted English as its most widely taught foreign language because English is now an international language.

But, why has English rather than any other language become an international language? This seems to be a rather complex question. But if we consider this issue from the following four aspects, we may gain a greater understanding.

a) Historically: The English language, as well as the English culture, spread to Asia, Africa, North America and Australia with the establishment of the British colonies in those four continents.

b) Geographically: In this world, there are six continents inhabited by human beings, of which Europe is greatly influenced by the English language, while North America and Australia are completely occupied by the English language. Besides, in many countries in Asia and Africa, English is an official language. No other language has such a power.

c) Politically: After the Second World War, both the United States and Britain, the two major English speaking countries, were amongst the most politically influential countries.

d) Economically: Since the end of the World War II, the United States has been the most powerful country economically with the most advanced science and technology. As a result, it has also become the most powerful country militarily. Britain has been amongst the most developed countries in the world. The economic position of both USA and Britain leads to the rapid development of media that have wide influence in the world.

When answering the question 'Why is the world becoming smaller and smaller?' we usually say that it is because of the rapid development of communication and transportation. In fact, we seem to ignore another very important factor: the globalisation of English. It is English that most readily bridges different cultures in the world. If the central contradiction of the 21st Century is between global networks and local identities, English is a tool of both (Warschauer, 2000). Most people speak English not because English is their native language, but because English is the language they could share. That English has become an international language is the result of natural development rather than deliberate creation. But given the development of the world, it is accidental for English to become an international language; it is inevitable that there will be an international language. That is to say, if English were not the international language, there would certainly be another language to become the international language.

The influence of the English speaking countries are indispensable to the fact that English has become an international language. Once English becomes an international language, however, neither the United States, nor Britain or any other English speaking country can control the language any more. Even native speakers may need to learn new dialects, not only to fully understand what people are saying in different parts of the world, but also to communicate effectively in international settings where the use of North American, British, or Australian colloquialisms may be inappropriate (Warschauer, 2000). Cook (1991) says that an Englishman or an American has no more right to tell a Tanzanian how to teach English than does a Chinese. More and more people use English not because it is the native language of the United States or any other English speaking countries but because it is the most powerful language belonging to the whole world. Even if the United States switched to another language, it would not change the status of English being the international language. Crystal (1997) believed that if there were to be a major social change in Britain that affected the use of English there, it would not have any real effect on the world trend.

The globalisation of English has a deep influence in China, which is reflected in the following:

a) there are specialized journals on English learning and teaching as a second language;

b) China Central Television Channels 4 and 9 are specialized channels in English;

c) in large cities, the special program of English News and English teaching are broadcast at regular times;

d) English newspapers, such as China Daily and the 21st Century, are becoming more and more popular;

e) people can often read English advertisements either in newspapers or magazines;

f) every hotel involving foreign countries or nationals is required to have instructions written in English;

g) slogans and illustrations in English are often seen in schools; and

h) enterprises of joint-venture or exclusive foreign investment use English more often.

What is mentioned above consists of the English language environment in China. In such an environment, people who are good at English are never unemployed. One always will be thought of highly if one has very high competence in English. Usually, enterprises do not employ graduates without English Certificate CET 4. Other conditions being equal, if one has English Certificate CET 6 and the other person has CET 4, one will be a favourite.

Prospects of English teaching in China

Teachers of English

English teaching and learning is not only a matter of language. Nobody can learn a language well without knowing its culture. At present, the vast majority of teachers of English learn English in China, teach English in China and finish their postgraduate studies in China. In other words, they usually have not had the experience of living in English speaking countries. Such a situation is not favorable for the improvement of English teaching. In recognition, the Chinese government has begun to send its English teachers to English speaking countries for further training and experience Such a program is considered to be a very promising beginning. Although compared with the total number of English teachers in China those who have been sent out are still a very small number. When everything is taken into consideration, the key issue of sending teachers of English to English speaking countries only involves money. Therefore, it seems to be a simple rather than complex issue. Thus, with the financial improvement of China, the day when all the teachers of English have the experience of being immersed in English culture will soon come.

Decrease of classroom teaching and increase of teaching and learning in other ways

In the past, classroom teaching has been the only approach to language learning and teaching. However, in recent years, there have been some significant changes, especially with the appearance of multimedia software for teaching and learning. In China there is one set of English textbooks for second or third year college students and three major English textbooks for undergraduates. For each of these English textbooks, there is at least one set of multimedia software for teaching and learning. All of these software are very convenient to use. For example, if one wants to improve one's pronunciation, they simply press the button for reading aloud and can read after a native speaker who has beautiful pronunciation and intonation. If one wants to know the Chinese meaning of any part of the English text, they press the translation button and the Chinese version will appear on the computer screen. Immediately after finishing the exercises the program provides their correct and incorrect answers. With such software, the classroom is no longer the only place to learn English, students could learn English even at home. The appearance of such software could be regarded as the turning point of English teaching and learning in China, for it will lead to multi-modal ways of English teaching and learning.

Beijing Foreign Languages Studies is now undertaking a national project on distance education of English. In his introductory speech to the Third International Symposium on ELT, held in Beijing 2001, the professor heading the project suggested that the future of English teaching was through distance education. Instead of learning English in classrooms, most people will be able to finish their English studies at home.

English teaching environment

As far as the English teaching environment is concerned, while efforts from learners and teachers, schools and communities produce micro-pictures, policy making decides the macro-picture. It is very hopeful that China will join the World Trade Organisation in the near future. This means that there will be an increase in the use of English in China. It seems to be a necessity to improve the English learning environment to help learners access more English. Such a situation will influence policy-making in China to accelerate the development of English language environments so much so that the English language becomes a second rather than a foreign language. Actually, that English is taught from even Grade 3 in elementary schools provides an exciting symbol of that fundamental change. The more the Chinese government looks at English teaching from a strategic perspective, the sooner the day will come when English becomes a second language in China.

CONCLUSION

English teaching has been a career of several generations of teachers of English in China since 1978. Lakoff (2000, p. 41) says,

> Language is not "just words". It enables us to establish our selves, and ourselves, as individuals and as members of groups; it tells us how we are connected to one another, who has power and who doesn't.

To some extent, development and improvement in China depends on the quality of its English teaching. In the age of information technology where English is becoming an international language, Chinese people still use English as a foreign language to get information while other peoples use English as their native or second language to do so. This places China at a disadvantage in language. Consequently, English teaching in China is much more difficult than in those countries whose English is either native language or a second language. Cook (1991) points out that in societies where the second language is in actual use, second language learning in the world outside can be compared with second language learning inside the classroom. With the globalisation of the English language, it seems that English teaching in China is not only a matter of a foreign language teaching. It has a very close relationship with the development of the whole country. Robertson (1992, p.8) says that 'globalisation as a concept refers both to the compression of the world and the intensification of consciousness of the world as a whole'. If both the Chinese government and people could look at this issue from a strategic point of view, it could be seen that China still has a very long way to go. To only teach English is not enough; how to make the quality of English teaching comparable with the English taught in other countries is what really counts.

REFERENCES

CET Commission (1999). *What Does CET Test: Analysis of Students' Performance on CET Papers Band 4*. Shanghai: Shanghai Foreign Language Education Press.

Syllabus Revision Team (1991). *College English Teaching Syllabus*. Shanghai: Shanghai Foreign Language Press.

Cook, V. (1991). *Second Language Learning and Language Teaching*. London: Edward Arnold.

Crystal, D. (1997). *English as a global language.* Cambridge: Cambridge University Press.

Deng, X. P. (1977), (1978) and (1979). *Selections from Deng Xiaoping: 1975-1982.* Beijing: People's Education Press.

Hogan, T. (1996). Globalization: Experiences and explanations. In A. Kellehear (Ed.), *Social Self, Global Culture*. Oxford: Oxford University Press.

Hu, W. Z. (In Press). A matter of balance: Reflections on China's language policy in education. *The Third International Symposium on ELT,* Beijing China: Foreign Language Teaching and Research Press.

Lakoff, R. T. (2000). *The Language War*. Berkeley: University of California Press.

Mao, Z. D. (1949). On people's dictatorship. In Selections from Mao Zedong (1991 ed.). Beijing: People's Education Press.

Ministry of Education PR China (2000). *English Syllabus for Junior High School*. Beijing: People's Education Press.

Ministry of Education PR China (2000). *English Syllabus for Senior High School*. Beijing: People's Education Press.

Neave,G. (2000). Introductions in higher education. In G. R. Teasdale & Z. M. Rhea (Eds) *Local Knowledge and Wisdom in Higher Education.* Oxford, England: Pergamon.

Robertson, R. (1992). *Globalization*. London: Sage.

Warschauer, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly*, 34, 511-535.

Zai,X.J., Zheng, S. T. & Zhang, Z. J. (1999). *21st Century College English*. Beijing, Shanghai: Higher Education Press, Fudan University Press.

Student perceptions of the importance and value of a Graduate Quality framework in a tertiary environment

Vicki Feast
University of South Australia *vicki.feast@unisa.edu.au*

The importance and significance of the attainment of graduate or generic skills by tertiary students is increasingly being recognised by tertiary institutions, employers, students and governments both nationally and internationally. This study examines the importance and value that students place on a Graduate Quality framework developed at the University of South Australia. The study involved administering a questionnaire to a group of 161 tertiary business students. Analysis of the questionnaire established a number of quantitative findings. Subsequently six students were interviewed in order to follow up the quantitative findings and to further investigate and establish reasons for these findings.

This research found that students place a good deal of importance on having a Graduate Quality approach to tertiary education and that students tended to undervalue their own achievement of Graduate Qualities but valued help from their courses in developing these Qualities. Some gender, age and career differences were found.

Graduate Qualities, Generic abilities, Tertiary education, Skills, Competencies

INTRODUCTION

Graduate Qualities can be described as those skills, which are useful in a range of work and life situations. They are intellectual, aptitudinal or interpersonal skills rather than attitudinal skills and are known by a variety of alternative names such as competencies and abilities. They focus on student outputs rather than inputs such as staff qualifications, the latter being the more common emphasis in assessing the performance of most universities. Students can achieve these skills as a consequence of their university studies but also outside university through work experience, community work and as a result of life experiences.

The type of skills classified as Graduate Qualities vary considerably across discipline areas, universities and nations as does the description, interpretation, achievement standard and meaning of each of the skills. For example, team work skills are often included in this definition but what is meant by team work skills, how they are measured and described and the standard of team work skills expected at various year levels of tertiary study vary considerably. Anticipated outcomes of developing these skills at university are also subject to debate. Expected benefits can extend from a micro level in terms of improving an individual student's prospects of employment to the macro level of improving a nation's ability to restructure the economy through microeconomic reform enabled by a more flexible and skilled workforce.

Various stakeholders are interested in the concept of Graduate Qualities, which have attracted increasing levels of scrutiny and debate in the last decade in Australia and last 25 years in the United States of America, Canada and the United Kingdom. Students enjoy the benefits of transitioning to an increasingly uncertain and ever changeable and flexible world of work

through the acquisition of these skills. Employers value students increased preparedness for work. Careers counsellors and graduate recruiters, both aware of these developments, are attuned to helping students with these skills and to employers more actively recruiting from institutions who assist students in skills attainment. Universities are under more external pressure to take a more significant role in producing skilled graduates for the workforce and the national good. Universities are also increasingly aware that they may be able to attain a competitive advantage in the market place over other universities who focus on content development alone. Governments who are interested in improving the attainment of national economic and social goals also value skill development and enhancement.

Implications of the increasing focus on skill development for universities include changes to pedagogy and assessment methods which are necessary to make explicit the development and attainment of Graduate Qualities. Implicit assumptions that students will acquire these skills by default during their education are slowly being replaced by an expectation that these skills will be made explicit in teaching methodologies and in assessment. Additionally the role for universities is seen now to extend to a more extensive preparation and transitioning of students to the world of work through programs such as Recording of Achievement (RA) projects. RA projects, which originated in the United Kingdom and United States of America, are now beginning to appear in Australian universities. They involve academic staff assisting students to document their attainment of these skills and develop a comprehensive portfolio of skill development. These portfolios play an important role in contributing to the construction of employment applications by students.

Within this climate of a growing awareness of the importance and value of skill development in tertiary education, it is becoming increasingly important that student views are researched, understood and accounted for. Past research has found that positive student perceptions play a significant role in the success of such developments. Some significant gender, age, and career differences in these perceptions are also noted in past studies. Research has also found that students feel that universities have placed too great a focus on learning content with too little emphasis on the importance of skill development in courses.

This paper is set in context at one Australian University, the University of South Australia (UniSA). UniSA has developed a framework of seven Graduate Qualities, which guide teaching and assessment developments in all courses. They were defined and initially developed by Senior Management of the university in 1995 and were discussed, debated and fine-tuned by members of academic staff in the following year. This fine-tuning included the writing of a number of indicators of each Quality to enhance their interpretation by staff and students. Since 1997 these Qualities have been progressively integrated into the teaching of subjects and courses. Considerable amounts of resources and professional development activities have been devoted to assist staff in understanding and developing them.

Within this Graduate Quality framework at UniSA a RA project was conducted in 2000 in the School of International Business. This project encouraged students to articulate their experiences and achievements in line with the UniSA Graduate Quality framework and to document them in a portfolio. The intention was that this portfolio would be used by students in their employment applications and would differentiate University of South Australia students from other graduates in the graduate labour market.

With the view to improving the understanding of the perceptions of students about the importance and value of a Graduate Quality framework, this study sought the insights of students in this RA project. A questionnaire was administered to 161 of the 345 students participating in this project seeking an answer to three specific questions about student perceptions regarding this framework. In line with previous research, statistical analyses of the questionnaire data were also used to investigate gender, age, and career differences about student perceptions of the importance of the Graduate Quality framework. Thus four key research questions were posed:

1. How important are Graduate Qualities to students in their quest for selection for employment in their chosen careers upon graduation?
2. How well do students perceive that their courses at the UniSA have equipped them to develop these Graduate Qualities?
3. How do students rate themselves in their individual/personal development of Graduate Qualities?
4. Are there significant differences in the relative importance of student perceptions of Graduate Qualities amongst gender, age and career groups?

METHODOLOGY

Two distinct methodologies were used in this study. First, a self-completion questionnaire administered in a face-to-face RA student meeting was designed to elicit student responses to the first three questions above. Question four was addressed using statistical procedures based upon the questionnaire data obtained.

The questionnaire methodology was chosen because it was a cost effective and time efficient way of collecting a relatively large amount of data and enabled standardisation of questions about student perceptions. This choice accords with the advice of Allan and Skinner (1991, p.173) who stated that in a choice of research methods one must " always keep an eye on what is feasible within the available resources of time, money and accessibility of resource material". Questionnaire methodology also enabled the researcher to ask questions of students, which were very close adaptations of research questions one to three. In this way, the research questions could be investigated first hand rather than from second hand data.

Second, six qualitative interviews were conducted to seek in depth explanations for the student perceptions gained in the questionnaire for research questions one through to three. Interviewing was chosen because of its flexibility and because it had the advantage of being able to "probe for more detail, seek reflective replies and ask questions which are complex" (Allan et al, 1991, p.228). In this way it enabled the quantitative data collected to be further analysed using a complementary yet rich, in-depth source of qualitative data.

Questionnaire Design and Description

A questionnaire (an excerpt of which is attached as Appendix one) was designed to provide insights into the above four questions. In this questionnaire close adaptations of the first three questions were posed for each of the seven UniSA Graduate Qualities. Students rated their perceptions about the first three questions for each Quality using a Likert scale of one through to five. Space was also made available at the end of each question for comments.

Question four was incorporated into the questionnaire design with the inclusion of a section at the beginning of the questionnaire asking students to state their gender, age career and course. Significance testing allowed research question four to be addressed for the 'importance' issue only. It was decided that the significance of gender, age, career and course groups for the 'course help' and 'achievement' questions was beyond the scope of the study and of lesser interest than for the 'importance' question.

One further decision in the make up of the questionnaire was made to facilitate this significance testing. Graduate Quality 4 was divided into 4a (working independently) and 4b (working collaboratively) because of the natural difference in the meaning of these two Qualities. Also Graduate Quality 5 was divided into 5a (committed to ethical action) and 5b (committed to social responsibility) for the same reason.

Interviews

Qualitative interviews were conducted to further investigate reasons for the findings with respect to the rankings of mean scores for questions one to three. At the end of the questionnaire students were asked to volunteer to be part of further longitudinal case study research. A sample of six of these volunteers was interviewed in order to gather more in-depth qualitative data. Seven interview questions (attached in Appendix two) were asked, based on the quantitative findings regarding student perceptions obtained in the questionnaire. The aim of the interviews was to further investigate the motivations and reasons for the quantitative data on student perceptions. These interviews were analysed and significantly added to richness of the findings of this study.

FINDINGS, ANALYSIS AND DISCUSSION

In this study 161 students, 75 males and 86 females completed the questionnaire from a total of 345 students. Student ages ranged from 18 to 43, 70 students were aged below 20, 80 students were aged between 20 and 30 and 11 students over 30.

Of the 138 students who specified their career choice these students represented 15 different careers. As is to be expected for students enrolled in a Business and Enterprise Division, they were mainly traditional business careers including accountant (32 students), banking (15 students), management (15 students) and tourism and hospitality (15 students). Other careers included sport, engineering and defence/police forces.

Each of the four research questions is considered in order. Findings for each research question are discussed followed by analysis and discussion of each question.

Research question one: How important are Graduate Qualities to students in their quest for selection for employment in their chosen careers upon graduation?

Table 1 and Figure 1 show rankings of the mean scores for each Graduate Quality by 'importance'.

Table 1. Ranking of "importance" means

Rank	Graduate quality no.	Mean Score (available range from 1-5)
1st	6 (Communication)	4.66
2nd	3 (Problem solving)	4.43
3rd	4b (Works collaboratively)	4.42
4th	2 (Lifelong learner)	4.41
5th	1 (Body of knowledge)	4.39
6th	4a (Works autonomously)	4.29
7th	7 (International perspective)	4.16
8th	5a (Ethical action)	4.02
9th	5b (Social responsibility)	3.94
Overall mean		4.3

Table 1 and Figure 1 show that the minimum score for importance of the Graduate Qualities was 3.94 out of a maximum score of five indicating that students value them highly in their quest for employment in their chosen careers.

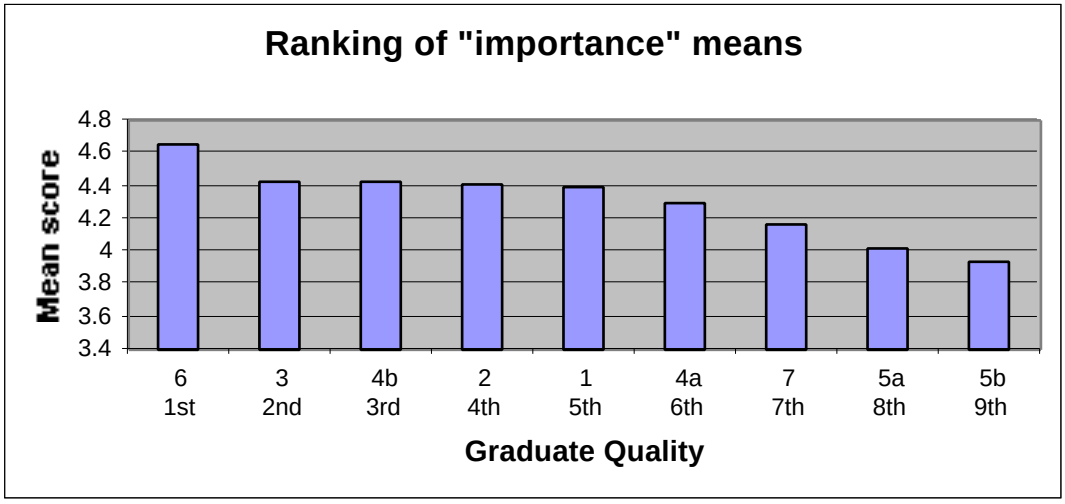


Figure 1. Ranking of "importance" means

An overall average of 4.3 placed the mean score between 'important' and 'very important' scales and a maximum mean score of 4.66 which is quite close to 'very important' also provided clear signals about high student perceptions of the importance of these qualities. The most important in the rankings in order from most important down were Graduate Qualities 6 (communicates effectively), 3 (effective problem solver) and 4b (works collaboratively). The least important in order from least important up were Graduate Qualities 5b (committed to social responsibility), 5a (committed to ethical action) and 7 (demonstrates an international perspective).

From the data shown in Table 1 and Figure 1 it is evident that students perceived Graduate Qualities to be important to their employment prospects. A high average overall mean score of 4.3 and high minimum (3.94) and maximum mean scores (4.66) gave a clear picture of student perception of 'importance'. In a survey of graduands in 1989, Smith, Wolstencroft & Southern (1989, p.28) found that "the striking feature is the high importance attached to most of the skills". Nabi et al (1998, p.5) found that all of the attributes surveyed in their study "were rated at least 'useful' ". The findings of this paper concur with this past research.

In the interviews students were asked the following question. Why do you think that many students in the RA trial rated the Graduate Qualities framework as important to their career development? Their replies are discussed next in an attempt to further probe reasons for importance of the Graduate Qualities.

All six students interviewed could see the importance and thus the value of Graduate Qualities in helping them find work. In one student's words they are also "fundamental to being in the workforce". Student one, a student most positive about the Graduate Quality framework, focused on it as providing a 'foundation' and a 'tool' to help him find work. He said many times that it improved his resume and would help him in an interview. Student six thought that the principles applied to all jobs "across the board", including part-time work in a supermarket and was not limited to careers. Student three thought they helped her "find out more things about yourself than you would have ever realised" and "helps you realise you're probably more worthwhile than you think". She felt that in a job application she was now better prepared to address all of the essential and desired criteria rather than giving up on some as she had done in the past. Student four and Student two both thought that it was more 'external' than 'internal', that is, the value was more in the documenting of Graduate Qualities than in the real changes Graduate Quality made to a student as a person.

Student four was also a little more cynical than the others in that he thought that the way Graduate Qualities were promoted in the RA program and the promotional materials had a large influence of students' positive attitude towards this question. It must be acknowledged

that the RA project made all participants more aware of the benefits of the Graduate Quality framework and probably more disposed to answer this question in a positive manner.

Research question two: How do students rate themselves in their individual/personal development of Graduate Qualities?

Table 2 and Figure 2 rank the mean scores for each Graduate Quality by student perceptions of 'achievement'.

Table 2. Ranking of "achievement" means

Rank	Graduate Quality no.	Mean Score (available range from 1-5)
1st	4a (Works autonomously)	3.86
2nd	4b (Works collaboratively)	3.81
3rd	6 (Communications)	3.76
4th	5a (Ethical action)	3.58
5th	2 (Lifelong learner)	3.54
6th	3 (Problem solver)	3.5
7th	5b (Social responsibility)	3.46
8th	1 (Body of Knowledge)	3.37
9th	7 (International perspective)	3.3
Overall mean		3.58

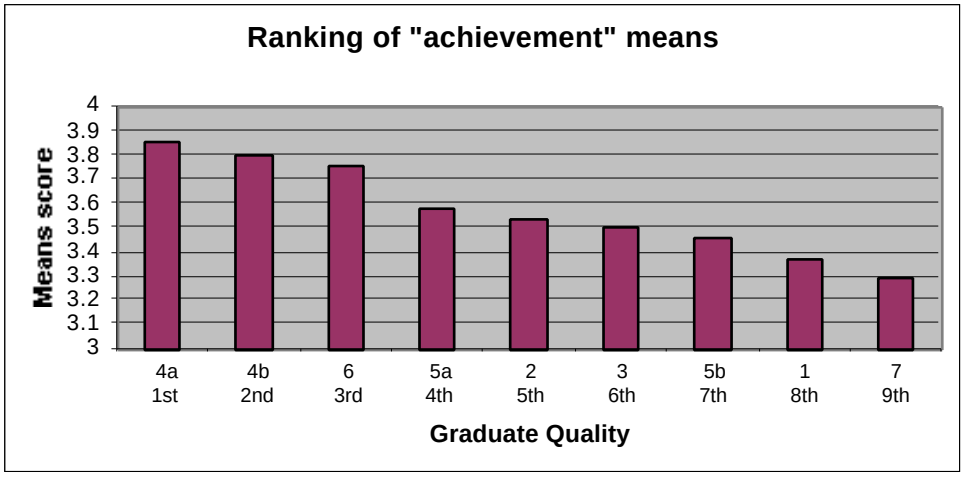


Figure 2. Ranking of "achievement" means

Table 2 and Figure 2 show that the minimum mean score for 'achievement' of the Graduate Qualities was 3.3 out of a maximum score of five indicating that some students rate their personal levels of achievement quite poorly at not much above 'just achieved'. An overall average of 3.58 which placed the mean score between 'just achieved' and 'quite a high level of achievement' and a maximum mean score of 3.86 which is very close to 'quite a high level of achievement' provided a slightly more encouraging picture about student perceptions of their achievements. However, some allowance should be made for possible upward bias in self-ratings gained in questionnaires of this type. The most important in the rankings in order from most important down were Graduate Qualities 4a (works autonomously), 4b (works collaboratively) and 6 (communicates effectively). The least important in order from least important up were Graduate Qualities 7 (demonstrates an international perspective), 1 (operates with a body of knowledge) and 5b (committed to social responsibility).

The data in tables two and Figure 2 indicated student ratings of 'achievement' were moderate with an overall mean score of 3.58, a minimum mean score of 3.3 and a maximum mean score of 3.86.

In the qualitative interviews students were asked the following question. Why did you rate yourself relatively well/poorly (choose one) in your own personal achievement of Graduate Qualities? Their replies are discussed next in an attempt to further probe reasons for their moderate ratings of 'achievement' of the Graduate Qualities.

These moderate ratings seemed to reflect individual students' lack of confidence and self-esteem. Five of the six students felt they could do better and all students felt that their experiences at university only went part of the way in developing these qualities. Student one particularly lacked confidence and blamed his poor English skills and hinted that his ability is not that high although he tries hard and is a 'doer'. Student four felt that a lack of awareness of Graduate Qualities throughout his course influenced his middle ranking. Student six felt she did not put enough effort into her studies to do really well especially in first year. However, she rated herself at 4 _ for group work because of the practice with this over the years and the good experiences she had working in a group this year. Thus grades and achievement in assessment seemed to influence this rating to some degree. Student five, the only student who rated herself highly felt that she bought many of the skills with her and that they were not promoted or explicitly taught well enough in the courses.

Research question three: How well do students perceive that their courses at UniSA have equipped them to develop these Graduate Qualities?

Table 3 and Figure 3 rank the mean scores for each Graduate Quality by student perceptions of 'level of help' from the course.

Table 3. Ranking of "help from the course" means

Rank	Graduate Quality no.	Mean Score (available range from 1-5)
1st	1 (Body of Knowledge)	4.01
2nd	4b (Works collaboratively)	3.92
3rd	2 (Lifelong learner)	3.83
4th	6 (Communications)	3.81
5th	3 (Problem solver)	3.71
6th	4a (Works autonomously)	3.69
7th	7 (International perspective)	3.64
8th	5a (Ethical action)	3.36
9th	5b (Social responsibility)	3.34
Overall mean		3.7

Table 3 and Figure 3 show that the minimum score for 'level of help' was 3.34 out of a maximum score of five indicating that students regard the course as being at least 'a reasonable amount of help' in developing Graduate Qualities. A maximum mean score of 4.01 indicated that student perceptions are just above a 'lot of help'. An overall average of 3.7, which places the mean score between 'a reasonable amount of help' and 'a lot of help', also provided clear indications about the level of assistance given to students by their course to develop Graduate Qualities. The most help in the rankings in order from most important down were Graduate Qualities 1 (operates with a body of knowledge), 4b (works collaboratively) and 2 (is prepared for lifelong learning). The least help in order from least help up were Graduate

Qualities 5b (committed to social responsibility), 5a (committed to ethical action) and 7 (demonstrates an international perspective).

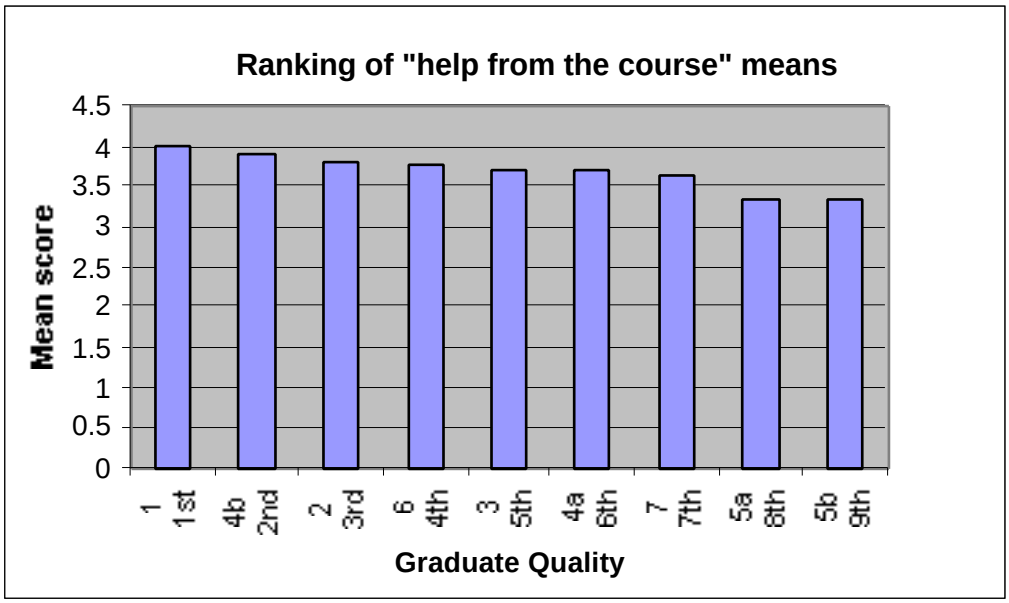


Figure 3. Ranking of "help from the course" means

Findings from Table 3 and Figure 3 indicated that students considered they gained a moderate degree of 'help from their course' in developing Graduate Qualities with an overall mean score of 3.7, a minimum mean score of 3.34 and a maximum mean score of 4.01.

In the qualitative interviews students were asked the following two questions about the course help. Their replies are discussed next in an attempt to further probe reasons for their belief that the course only provided a moderate amount of help in developing Graduate Qualities. Question three asked 'Why do you think that many students in the RA trial believed that their course helped them develop Graduate Qualities'?

Four of the six students didn't feel that the course helped them develop Graduate Quality a great deal, despite the average rating of 3.7 out of five in the survey. They appeared to be saying that their achievement of Graduate Qualities was more an incidental part of studying and they weren't fully aware of the framework. Student four thought UniSA gave students a 'good start' but a lot more could be achieved in the workforce. This perception of 'help from the course' varied amongst the students interviewed and was dependent on the individual Graduate Qualities. For example, Graduate Quality 1 mentioned quite strongly by Student two. However, Student two also thought "students learn more on the job in two weeks than in their whole course" questioned even the attainment of this Graduate Quality being due to the course. Student five thought that problem solving was focused on in assignments and that group work had some emphasis but that lifelong learning and ethical action (except in one module in one subject) were not taught in any deliberate way.

The emphasis on Graduate Quality 1 is to be expected as a traditional emphasis in universities has been on learning a body of knowledge. As these were final year students and the Graduate Quality framework is relatively new at UniSA, it is also probably to be expected that the influence of this framework is only just beginning to explicitly impact on students. Student two did mention that his course is "rich in communication skills" and that there was some help with gaining an international perspective although his overseas travels were the biggest contribution to him gaining this skill.

In contrast, Student six and Student three were confident that some of the subjects helped them achieve some of the Graduate Qualities. Student six mentioned the content of two subjects which helped her gain an international perspective (Graduate Quality 7) because she learnt

about the economies of other countries. She thought Graduate Quality 1 was gained through textbook learning. She also mentioned that assignments presented problems to be solved in themselves so that helped her critical thinking (Graduate Quality 3) to "answer it relevantly as well as concisely in the best way". Working in groups and autonomously was mentioned by both students as assisting Graduate Quality 4, and communicating effectively (Graduate Quality 6) was promoted through written assignments and oral presentations. Student six also added that lifelong learning (Graduate Quality 2) was gained through her personal development at university and that plagiarism information and discussions helped her deal with Graduate Quality 5 (ethical action).

Question four of the interviews asked students 'What could have been done to improve this development in your own subjects/courses? – consider teaching methods, assessment etc'

A common theme (discussed by five of the six students) was that students wanted more emphasis by academic staff in lectures and tutorials on the framework and how it can be applied to each subject. These five students felt that it is necessary that academic staff keep the Graduate Quality framework at the forefront of students' attention. Student two felt that more emphasis on putting students in more realistic and unpredictable situations so they could show their initiative would be helpful although he acknowledged that resources limit this course of action. Student four felt that incorporating work experience into courses or making courses more practical would help. Student six felt Graduate Qualities should be included more in study guides and newsletters and made more explicit in assignment requirements. Student three thought Graduate Qualities should be mentioned "right from the word go" and right throughout the semester, making it "natural for people to talk about it". Student five thought there should be more emphasis on Graduate Quality 5, ethical issues and social responsibility.

Comparison of the three research questions

Table 4 indicates a summary of the findings for each of the above three research questions for comparison purposes.

The first row of Table 4 shows the ranges of each mean score and the highest and lowest mean scores, for the nine Quality categories for each of the three questions. It shows for example that for the 'importance' question the minimum mean score was 3.94 (for Graduate Quality 5b) and the maximum mean score was 4.66 (for Graduate Quality 6) with a range of mean scores amongst the Qualities of 0.72. The second row shows the overall mean score for all Qualities for each of the three questions. For example, the overall mean score for all of the Qualities for 'importance' was 4.3. The third row shows the Qualities with the highest mean score in order from the top down for each question. For example for the 'importance' question, Graduate Quality 6 (communicates effectively) had the highest mean score followed by Graduate Quality 3 (effective problem solver) and Graduate Quality 4b (works collaboratively). The fourth row shows the Qualities with the lowest mean score from the lowest up for each question. For example for the 'importance' question, Graduate Quality 5b (committed to social responsibility) had the lowest score followed by Graduate Quality 5a (committed to ethical action) and Graduate Quality 7 (demonstrates an international perspective).

From an analysis of Table 4 it was apparent that students rate the importance of the Graduate Qualities (4.3 mean score) higher than their help from the course (3.7 mean score) with individual achievements (3.58 mean score) ranking last of the three questions. In short, the students felt that Graduate Qualities were important but the University was not as helpful in getting them to achieve these qualities nor do they think their achievements of the qualities, deriving from both university and external sources, were particularly high. This finding concurs with other research such as Nabi & Bagley (1998, p.4) who found that students rate their achievements below the level of importance. This may also mean that the UniSA still has

some way to go with improving on how courses help students to bridge the gap between importance and help given by the course.

Table 4. Comparison of range and means for three questions

Range of each mean	Q1 (importance) – 4.66 – 3.94 (0.72)
	Q2 (achievement) – 3.86 – 3.3 (0.56)
	Q3 (help from course) – 4.01 – 3.34 (0.67)
Overall means	Q1 (importance) – 4.3
	Q2 (achievement) – 3.58
	Q3 (help from course) – 3.7
Qualities with the highest scores (highest three in order from top down)	Q1 (importance) - 6, 3, 4b
	Q2 (achievement) – 4a, 4b, 6
	Q3 (help from course) – 1, 4b, 2
Qualities with the lowest scores (lowest three in order from bottom up)	Q1 (importance) – 5b, 5a, 7
	Q2 (achievement) – 7, 1, 5b
	Q3 (help from course) – 5b, 5a, 7

Generally Graduate Quality 4, especially 4b featured in the highest three ratings for all three questions and Graduate Quality 5, especially 5b rated in the lowest three ratings. Thus it seems that students rated working collaboratively as a key skill and a commitment to social responsibility as a less likely achievement regardless of the questions asked.

The relatively high ratings of working collaboratively and the relatively low ratings of commitment to social responsibility are not the intention of the positioning of the Qualities by UniSA. Each Quality is treated as equally important in the framework. It may be that these perceptions of importance are peculiar to business students and that other students may have different perceptions. Further study to investigate these issues is strongly recommended.

Graduate Quality 7 (an international perspective) featured in the lowest scores, a fact which may raise concerns in a School of International Business in which most of these students study. Graduate Quality 6 (ability to communicate) featured in two of the top three ratings (importance and achievement) but not in level of help from the course. This may mean that more emphasis in courses on teaching and assessing effective communication skills is required and that students felt these skills, to the extent they were achieved, were achieved despite, rather than because of, the help from their course.

In the interviews students were asked the following two comparative questions about the first three research questions. Their replies are discussed next in an attempt to further probe reasons for these beliefs. Question six of the qualitative interviews asked ‘Can you give any reasons why students in the RA trial rated the importance of the Graduate Qualities higher than their own development and the help the course gave them’?

Reasons for rating ‘importance’ above ‘help’ and ‘ achievement’ as mentioned by the six students included:

- Many experiences/achievements were non university related
- Students tend to underrate themselves (two students)
- Four students commented that they can always improve so it is difficult to give themselves high ratings especially if they haven’t done well in their course

- Two students felt that the importance of Graduate Qualities is plain and very easy to see but it is harder to see how the course helps and how you personally rate as a student
- Students rated their achievements over a broader spectrum (one through to five) than they do the importance
- A lack of awareness of how each assignment addresses each quality and thus the help students were receiving is not made explicit
- A lack of explicit 'teaching' about the qualities

Question seven asked 'Can you give any reasons why Graduate Quality 4 (especially Graduate Quality 4b, works collaboratively) rated relatively well and Graduate Quality 5 (especially Graduate Quality 5b, social responsibility) rated relatively poorly in the RA trial survey'?

Students felt that Graduate quality 4 was emphasised in courses a lot, especially Graduate Quality 4a (working independently), but also Graduate Quality 4b (working in groups to complete assignments). These Qualities were almost universally seen as important or even essential in the workforce and emphasised in job advertisements, particularly Graduate Quality 4b (working collaboratively).

All students agreed that Graduate Quality 5 (ethical action and social responsibility) was not emphasised in many subjects and may not be viewed as important as making money in a business related course. Student two saw ethical action as a "side issue" which businesses can only consider when they are "comfortable or have made it". Student six saw that ethical action was more related to the external environment and developed more through her part-time work and was a personal thing related to her conscience. She felt that university only supported her to do the right thing in a task-orientated way such as avoiding plagiarism and being culturally sensitive in her use of language. Student six also felt that Graduate Quality 5 was also not emphasised in job advertisements and interviews so students saw it as less important. She also commented that ethical action is culturally specific. What is ethical action to one culture is not to another, a prime example being whaling within Japanese culture being accepted but not so within Australian culture.

Student two also indicated that the order of the questions in the questionnaire might have influenced the ratings and that students may have given lower ratings as they worked through the questionnaire and became tired of it.

Research question four: Are there significant differences in the relative importance of Graduate Qualities amongst gender, age and career groups?

Gender

Findings with respect to the significance of gender are shown in Table 5. Var one through to seven refer to individual Graduate Qualities. Student's t –test was used to test for the significance of gender in rating the importance of Graduate Qualities.

Table 5. Significance of mean scores for 'importance' for each Graduate Quality based on gender.

Gender	N	Var1	Var2	Var3	Var4a	Var4b	Var5a	Var5b	Var6	Var7
Female	78	4.33	4.43	4.5	4.33	4.41	4.04	4.01	4.66	4.18
Male	66	4.45	4.38	4.34	4.25	4.43	4	3.91	4.66	4.13
	P-value:	0.27	0.65	0.28						0.73
	Sign?	No	No	No	No	No	No	No	No	No

As Table 5 shows, this study found no significant differences between the perceptions of males and females for 'importance' for any of the Graduate Qualities. Although females were found to have slightly higher mean perception ratings for six of the nine Graduate Quality categories compared to males, none of these differences were significant at the 0.05 level.

The findings of this study which showed no significant differences in male and female perceptions about the importance of Graduate Qualities are contrary to some other research. Nabi et al (1989, p.5) found that " males and female graduates had different views about the importance and quality of their skills" and that "females rated most of the skills as more important than males" although "males tended to rate the quality of their own communication and problem-solving skills more highly than females".

Table 6 summarises the findings of this study with respect to age and using analysis of variance significance tests. As previously reported, students whose age ranged from 18 to 43 were grouped into three groups (less than 20, between 20 and 30 and over 30) for the purposes of this analysis.

Table 6. Significance of mean scores for 'importance' for each Graduate Quality based on age.

	N	**1**	**2**	**3**	**4A**	**4B**	**5A**	**5B**	**6**	**7**
18 - 20	70	4.38	4.33	4.31	4.13	4.37	3.83	3.81	4.62	4.01
21 - 29	80	4.36	4.43	4.52	4.36	4.42	4.09	4.03	4.67	4.30
>= 30	11	4.64	4.73	4.56	4.82	4.7	4.64	4.45	4.82	4.18
P-value		0.421	0.153	0.343	0.009	0.437	0.009	0.037	0.593	0.128
Sign?		No	No	No	Yes	No	Yes	Yes	No	No

Age

These findings show that for Graduate Qualities 4a (works autonomously), 5a (committed to ethical action) and 5b (committed to social responsibility) older students (over 30) were significantly more likely to rate these skills more highly (P-value less than 0.05). For all other Qualities there was no significant difference in student perceptions of importance based upon age.

Little previous research has been found about the age characteristic. This study found that older students feel that Graduate Qualities 4a, 5a and 5b were more important to them than younger students. Given the low mean perception ratings in general for Graduate Qualities 5a and 5b, it may be tentatively suggested that the importance of these skills only become evident to students as they gain from their life experiences. Anecdotally, it is the author's experience that older students prefer to work independently and younger students prefer to work in groups. These findings give some weight to this experience. For the remaining Graduate Qualities there were no differences in student perceptions based on age.

Career

Because of insufficient numbers of students in some categories and the fact that students were mainly from business careers, the findings of this study are limited to examining significant differences between the perceived importance of the Graduate Qualities in six business careers. Of the 15 career categories, six careers (highlighted in Table 7) with 10 students or more were examined using student's t test. Table 7 summarises these findings. Note that var1 through to var7 refer to the Graduate Qualities of the same number.

A comparison of the scores for the highest and lowest mean score for 'importance' of each of Graduate Qualities (shown as var 1 onwards) in career development (question one in the survey) is shown in Table 7. In four of the six careers a significant difference was found between the two extremes (P-values of less than 0.05). For Tourism and Hospitality and

Banking careers no significant difference was found in the maximum and minimum mean 'importance' scores. This means that students in these two degrees did not see a difference in the importance of individual Graduate Qualities. For two careers, namely Human Resources and Property careers Graduate Quality 6 (communicates effectively) was found to be significantly more important to a students' career development than Graduate Quality 7 (demonstrates an international perspective). For Accounting careers Graduate Quality 6 was found to be significantly more important than Graduate Quality 5a (committed to ethical action) and in Management careers Graduate Quality 6 (communicates effectively) was significantly more important than Graduate Quality 5b (committed to social responsibility).

Table 7. Significance of mean scores for 'importance' for each Graduate Quality based on career

Career	N	Var1	Var2	Var3	Var 4a	Var 4b	Var 5a	Var 5b	Var 6	Var 7	Max	Min	P-value
Business/ Industry/ Analyst	5	4.2	4.8	4	3.8	4.6	4.6	4.4	4.6	4			
Banking/ shares	13	4.615	4.455	4.6	4.727	4.455	4	4.3	4.75	4.417	4.75	4	0.055
Sport	1	5	4	5	3	4	5	5	4	5			
Human Resources	12	4.583	4.167	4.333	4.538	4.538	3.923	3.846	4.846	3.769	4.846	3.769	0.0002
Accountant/ Financial Adviser-Planner	30	4.433	4.448	4.407	4.1	4.483	3.793	3.828	4.69	3.9	4.69	3.793	0.0001
Information Technologist	7	4.429	4.667	5	4.286	4.6	4.167	4.5	4.143	4.286			
Manager/ CEO	13	4.154	4.385	4.545	4.462	4.308	4.308	4	4.8	4.364	4.8	4	0.0031
Embassy/ Foreign Affairs	3	5	5	4	4	4	3.75	3.25	4.5	4			
Property Management/Valuation	10	4.3	4.15	4.5	4.2	4.5	4	4	4.6	3.95	4.6	3.95	0.019
Marketing	9	4.444	4.375	4.5	4.125	4.2	3.75	3.714	4.333	4.333			
Government	1	4	4	5	4	4	3	3	5	5			
Tourism/ Hospitality	15	4.467	4.533	4.667	4.5	4.633	4.429	4.25	4.667	4.467	4.667	4.25	0.22
Engineering	2	5	3.5	4	4.5	4.5	3.5	3	4.667	4.333			
Operations	1	4	4	5	5	5	4	4	5	4			
Defence/ Police Forces	2	4	4.5	4.5	4	4	4	4	5	4.5			

Smith et al (1989, pp. 28-29) found that jobs in manufacturing place more importance on creativity and critical thinking, in the commercial sector more importance was placed on teamwork, in the retail sector on communications and for law graduates problem solving was considered more important. They concluded that "graduate level jobs do make significant demands in terms of problem-solving skills, communication and working with others and that the exact nature of these demand appears to vary from one sector of employment to another" (1989, p. 30).

These findings concur to some extent with that of Smith et al (1989) in that there was some evidence of variance in the importance placed on individual skills in different careers. In particular communication skills were found to be more important than other skills in Human Resources, Property, Accounting and Management degrees. Commitment to social responsibility and ethical action were typically seen to be less important.

CONCLUSION

There is little doubt that students felt that the Graduate Quality framework developed at UniSA was important and contributed to their chances of employment in their chosen careers; and that their courses and subjects had contributed to some extent towards the development of those qualities. However, there is some need by the University to address implementation and student awareness issues in order to more fully capitalise on these findings.

Students rated the qualities chosen by the university as important to their quest for selection in their chosen careers but rated themselves less highly in their personal achievement of these Graduate Qualities. In the process of interviewing six selected students some underrating of achievement levels was reported by students to be caused by a lack of confidence and the low grades achieved by some students. Thus the findings of this paper also concur with other research, which found that graduates tended to highly rate the importance of these skills but that they rate the importance of the skills higher than their own ability in these skills.

Of the individual Graduate Qualities, Graduate Quality 4 (especially 4b, working collaboratively) and to a lesser extent Graduate Quality 6 (communication skills) were consistently given high ratings by students and Graduate Quality 5 (both 5a, ethical action and 5b, social responsibility) were generally at the bottom of the ratings for each question. Working independently and collaboratively and communicating effectively were thought to be much more relevant to students than a commitment to ethical action and social responsibility.

Some significant differences were found on the importance students placed on different skills by career, age and gender. In business careers such as management and marketing, students significantly valued effective communication skills more than their commitment to social responsibility and ethical action. On the other hand, significant gender differences were not found. Age was only a significant factor with three of the Graduate Qualities (4a, works independently, 5a, ethical action and 5b, social responsibility) where older students were found to place significantly more importance on these skills than younger students.

REFERENCES

Allan, G.and Skinner, C. (1991) *Handbook for Research Students in the Social Science*s, The Falmer Press, London.

Nabi, R. and Bagley, D. (1998) Graduates Perceptions of Transferable skills and Future Career Preparation in the UK, *Career Development International*, 3 (1).

Smith, D., Wolstencroft, T. and Southern, J. (1989) Personal Transferable Skills and the Job Demands on Graduates, *Journal of European Industrial Training*, 13(8), 25-31.

APPENDIX 1: SURVEY INSTRUMENT EXCERPT

Course:_____

Chosen Career:_____
(if you have more than one chosen career, list your first priority above)

Age_____ Gender_____

1) Rate the Graduate Qualities based on your perception of how important they are to you in getting you into your chosen career. (***Please rate your perception overall for each quality first and then rate each of the indicators in the same way ie wherever there is a star* a rating is required***).

1 = Of no importance 2 = Of little importance 3 = Some importance4 = Important 5 = Very important

Graduate Quality	Level of Importance
Body of knowledge	
2. Lifelong learning	
3. Effective problem solver	
4a. Ability to work autonomously	
4b. And collaboratively	
5a. Commitment to ethical action	
5b. And social responsibility	
6 Communicates effectively	
7 Demonstrate an International perspective	

Please provide any additional comments in this space

APPENDIX 2: INTERVIEW QUESTIONS

1. What single words or phrases would you use to describe your experiences with /feelings about the Graduate Quality framework during your studies at the University of South Australia?
2. Why do you think that many students in the RA trial rated the Graduate Quality framework as important to their career development?
3. Why do you think that many students in the RA trial believed that their course helped them to develop Graduate Qualities?
4. What could have been done to improve this development in your own subjects/ courses – consider teaching methods, assessment etc
5. Why did you rate yourself relatively well/poorly (choose one) in your own personal achievement of Graduate Qualities?
6. Can you give any reasons why students in the RA trial rated the importance of Graduate Qualities higher than their own development and the help the course gave them?
7. Can you give reasons why Graduate Quality 4 (especially Graduate Quality 4b) rated relatively well and Graduate Quality 5 (especially Graduate Quality 5b) rated relatively poorly in questions in the RA trial survey?

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

E-mail discussion and student learning outcomes: A case study

Ruth Geer
University of South Australia *ruth.geer@unisa.edu.au*

Dr. Alan Barnes
University of South Australia *alan.barnes@unisa.edu.au*

E-mail discussion is widely used in university courses and is being increasingly adopted in K-12 education. However making educational value from the massive amounts of e-mail transcripts that can be generated is a problematic task. It is important in an environment of constrained teaching resources that computer driven methods be developed to evaluate effective learning processes. A critical starting point is the relationship between messaging and student learning outcomes. This paper examines student learning outcomes and learning behaviours over three years of an education degree course involving e-mail discussion technology. It documents a clear relationship between the number of messages sent by students through e-mail discussion and the final grade for this course. Normative behavioural patterns that are evident from the transcript data can assist in ongoing educational monitoring.

e-mail, collaborative learning, interactive, Internet, evaluation.

INTRODUCTION

With the increased complexity of the online environment, educators are seeking ways to effectively utilise available technologies for quality teaching and learning. There are a number of factors, such as the type of activity and the assessment requirements that can impact on their effectiveness. Interaction and collaboration have been recognised as key ingredients in the learning process where students are not only active and supportive, but also interactive (Kirby, 1999; Hiltz, 1998; King & Doeffert, 1996). In collaborative learning environments students are actively engaged in supporting each other in the development of higher level reasoning strategies, critical thinking, hypothesis formation and reflection. Email discussion groups can be implemented to support collaborative and discursive interaction. A number of theoretical perspectives, with their foundations in cognitive developmental, behavioural and social interdependence theories (Johnson & Johnson, 1996) are guiding the design of courses and our understanding of how students learn.

Extensive research has shown the capability of computer-mediated communication to engender quality learning (Sringram & Geer, 2000; Newman et al., 1997; Gunawardena et al., 1997). More recently Geer & Barnes (2001) have developed rapid sampling techniques to identify metacognitive processing within textual discourse. A major thrust of that research was to provide an analysis of the content and interaction of the participants; and to identify indicators of cognitive and metacognitive learning that demonstrated the presence of quality learning. Such detailed analysis is very time consuming and not always practical during the duration of the course. With an overwhelming quantity of discourse being generated and the time that is required for a detailed analysis of the discourse, educators need to find ways that can readily give them a sense of the educational quality of the discourse so that strategic intervention can

take place if needed. Are there methods then, that can be used with email discussion technologies that can give an indication that students are learning through engagement with the course content?

The following *Model of Technology-Mediated Interaction* (Figure 1), in the shape of a pyramid, attempts to show a progression in the levels of interaction that show a shift from surface to deep learning (Ramsden, 1992). It has been developed in order to explain the relationship between types of interactivity that can occur in the online environment and the suitability of various technologies to support interaction and the associated learning outcomes. Each level denotes a type of interaction that might be expected in relation to specific teaching and learning approaches. There are many factors or drivers that can impact on the effectiveness of the interaction and assessment has been shown to be one of the factors that have a strong impact on learning outcomes. The intention of this model is to provide a conceptual framework, in which you can identify the entry point of the type of interaction required, utilising the various technologies that support such a level and ensuring that relevant learning outcomes are being achieved.

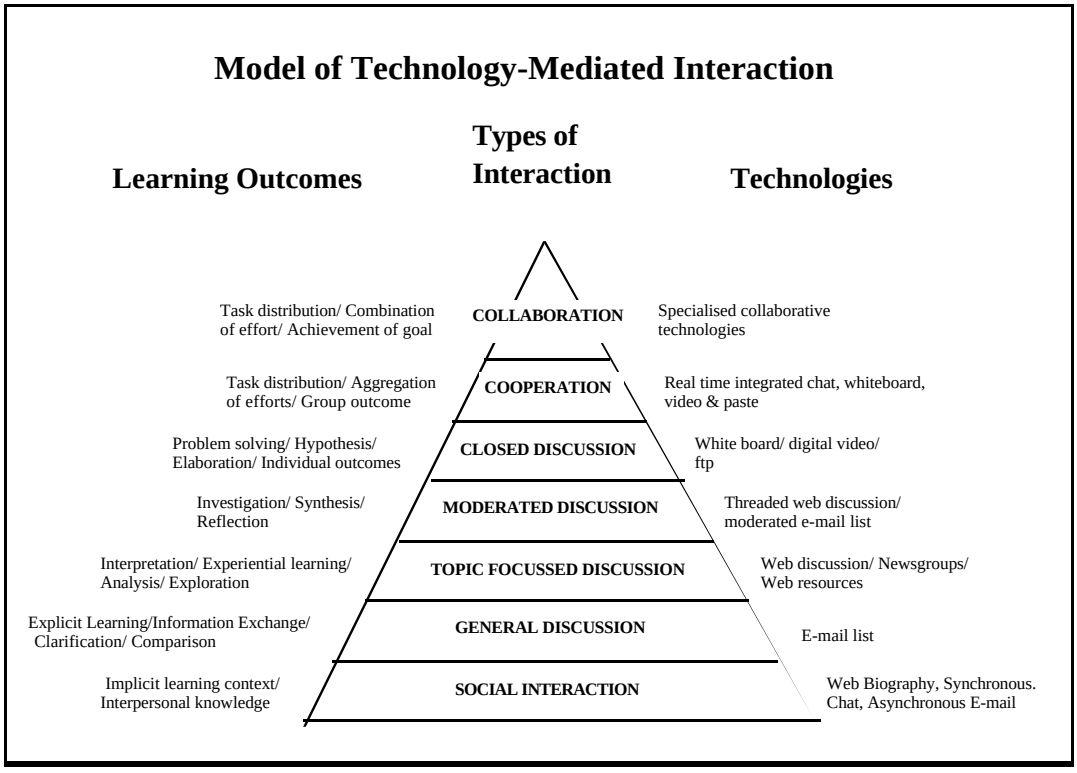


Figure 1. Model of Technology-Mediated Interaction indicates the hierarchies of interaction, technologies and learning outcomes, together with the influential drivers affecting learning outcomes.

BECOMING INFORMATION LITERATE: THE COURSE AND ITS CONTEXT

There have been a series of initiatives over the past few years in the course, Becoming Information Literate (BIL) to enhance learning outcomes, develop greater independence in students and improve the efficiency of teaching and delivery through the use of email discussion lists. The focus is on the development of effective computer skills, the integration of learning technologies into the curriculum and the opportunity to analyse underlying conceptual and educational issues. Email discussion lists have been closely embedded into the course assessment, as it was felt mandatory learner participation and interaction would increase the engagement with the course content through the development of the higher cognitive thinking

skills. It has provided the opportunity for student interaction to take place at any time and offered students the flexibility of time, place and pace in which to reflect and respond to topics and questions.

BIL is a core first year course within the Bachelor of Education program. Students are organised into tutorial groups of about 20 students and allocated to email discussion lists according to their tutorial groups. Approximately every three weeks students are assigned a topic for discussion via their email lists. Based on their readings in the area, students are expected to post reflective comments of 300-400 words in response to the topic to their assigned email discussion list. The responses are not marked as such, but they are a mandatory requirement of the course. Students did not have to give their response as a single message, but their feedback may span a number of messages. The majority however did provide an initial single message response to all the topics. As well as responding to the topics, students are assigned to a small collaborative group where they given responsibility for one of the discussion topics. Their task is to identify and summarise the key issues raised through the discussion lists and to present these in a face-to-face tutorial session. This encourages further interaction as the small group is challenged to critique the responses of their colleagues. It was felt that the instructional design would encourage increased interaction and collaboration. These strategies have been recognised as desirable ingredients in a learning community, as well as an essential quality in the social context. Based on the pyramid cited above it was also anticipated that students through their discussion would demonstrate processes of interpretation, experiential learning, analysis and exploration. It is recognized that associated learning benefits are closely tied to the assessment requirements.

E-MAIL TRANSCRIPTS AND STUDENT LEARNING BEHAVIOURS

Email transcripts from these discussions have been collected over the past three years resulting in approximately 2 megabytes each year of original e-mail discussion transcript. In the last two years a web archive has been made available providing increased flexibility. Students have also been able to use multiple e-mail aliases. Basic statistics on the e-mail traffic and length of messages are displayed in Table 1 below.

Table 1. Basic data on the number of students and the total number of e-mail messages and words sent to their respective discussion groups. Only original messages are counted, and only students who gained some marks in the course are included.

Year	Students	Messages		Words	
		Total	Mean	Total	Mean
1999	253	1150	4.55	460,023	1818.27
2000	186	1010	5.43	307,815	1654.92
2001	213	1386	6.51	402, 989	1891.97

There are no apparent differences in the number of messages or the message length by males and females in any of the years. However there does appear to be two different types of messages that can be identified by different characteristic lengths (Figure 2).

Students are encouraged to use the medium to increase their social comfort and to explore its possible potential. There are a considerable number of shorter messages, which can be associated with socialising, exchanging information, organising, and clarifying expectations. The other group of messages can be identified as the mandated contributions from each student. This group is almost symmetrically distributed about a mean of 425 words, which corresponds to assessment requirement of 400 words for each contribution. Both BIL 1999 and BIL 2001 show similar bi-modality to BIL2000.

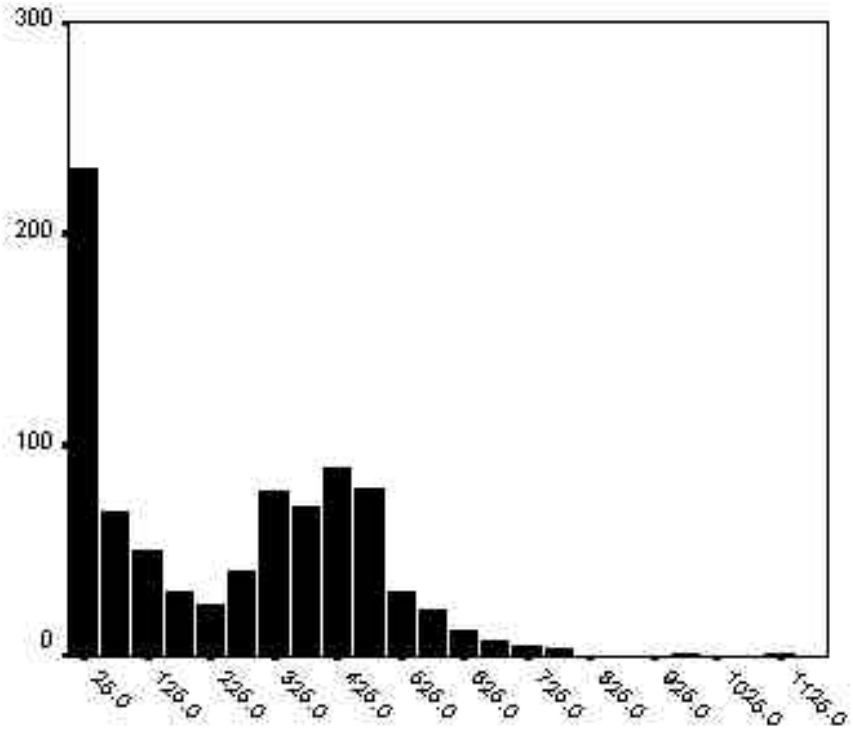


Figure 2. Frequency of message sizes in words for BIL 2000. Each step is 50 words

The impact of the assessment driver on transcripts is further demonstrated in Figure 3 where the peaks correspond to due dates for e-mail contributions. Such information could be used as a running analysis tool to check that responses have been sent on time. Lateness in submission impacts on the ability of the small group to summarise and critique the responses sent by peers.

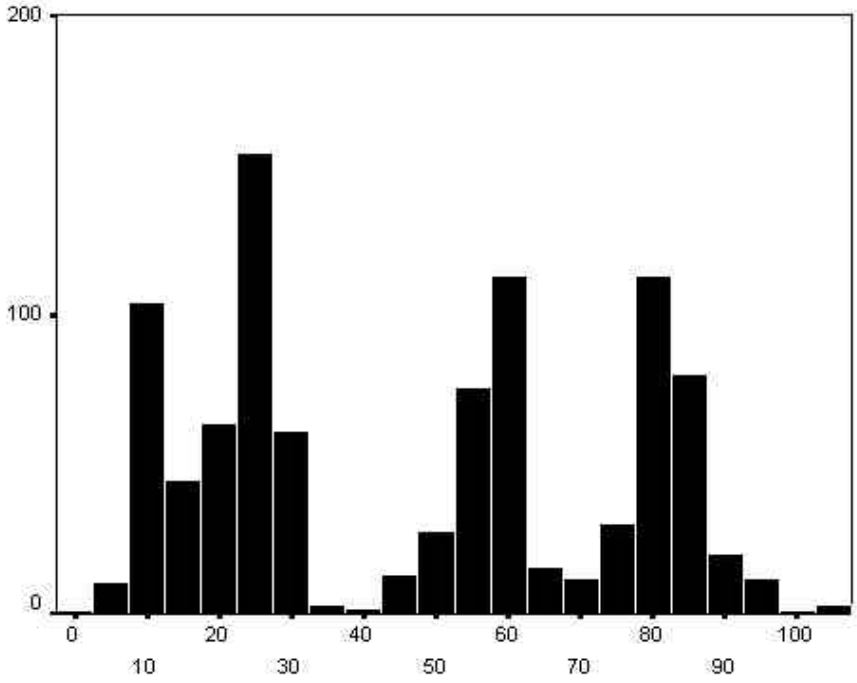


Figure 3. The graph charts the number of messages in a 5-day block starting from the beginning of the semester. The data is for BIL 2000.

The timing of e-mail submissions during the day illustrates the working behaviours of students. Approximately 75% of e-mail traffic is sent outside of formal tutorial times. This is a pleasing result as students are recognizing the need to focus on the planned activity for that session rather than being distracted by the need to submit their response during the tutorial session. Predominantly for this campus based course e-mail is sent between the hours of 9am and 2pm. However there is still considerable activity outside of the 9-5 working day with over 25% of e-mail messages being sent after hours. The following graph provides a clear picture of the email habits as messages sent during the tutorial sessions have been eliminated. There is a similar pattern of behaviours across each year. Over the three years the use of external mail providers has increased, almost doubling to some 10% of messages for BIL2001.

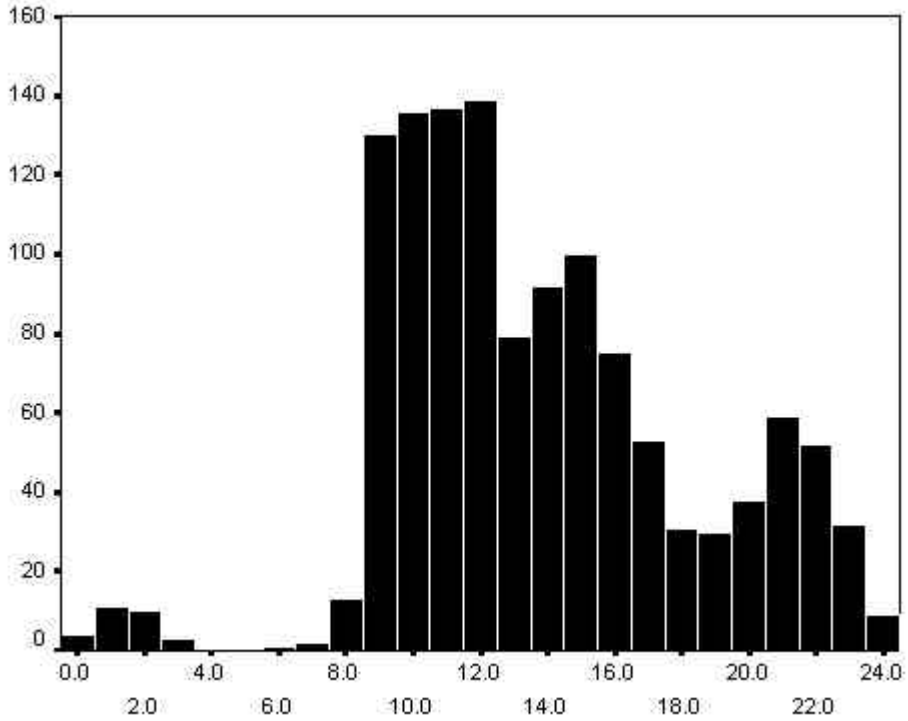


Figure 4. The graph shows the frequency of e-mail each hour around a 24-hour clock. This is for BIL 2001 and shows e-mail sent only outside of the standard tutorial times

At the time of analysis there appears to be no identifiable relationship between time of day and the size of the message. However with further analysis we may find that it reflects particular individual study habit behaviours.

E-MAIL TRANSCRIPTS AND STUDENT LEARNING OUTCOMES

In previous studies the authors and others have examined the discourse content of email transcripts for evidence of learning. More rapid techniques have been developed for identifying key individual and group metacognitive activity (Geer and Barnes, 2001), however these remain tedious and at best could only practically be used as sampling techniques to provide some assurance that learning is taking place. On the other hand a study of summative learning outcomes and their relationship with messaging can provide a picture of expected behaviour in a given course like BIL.

The following table shows the correlations between the summative grade for students in each of the BIL years and the frequency and size of messages.

Table 2. Pearson correlation coefficients for messages and message words with the total mark.

Year	Statistic	Messages	Message words
1999	Pearson Correlation Total Mark	0.24	0.19
	Sig. (2-tailed) N=253	0.0001	0.0022
2000	Pearson Correlation Total Mark	0.26	0.23
	Sig. (2-tailed) N=186	0.0004	0.001
2001	Pearson Correlation Total Mark	0.24	0.26
	Sig. (2-tailed) N=213	0.00045	0.0001

There appears to be a positive relationship between the overall learning outcomes for the BIL courses and messaging frequency. Generally better students are mailing more frequently and gaining more marks than others, even in the context that such mailing is not directly assessable. Figure 5 displays this relationship in graphic form for the BIL 2000 class.

Moreover the size of this relationship is fairly constant over the three years of the course, as demonstrated in the table of regression coefficients below. Student messaging behavior can thus provide an explicit measure of "time on task" and be used as a proportional indicator of student learning performance. It provides an extra formative tool beyond that of summative assessments to assist instructors in managing their students.

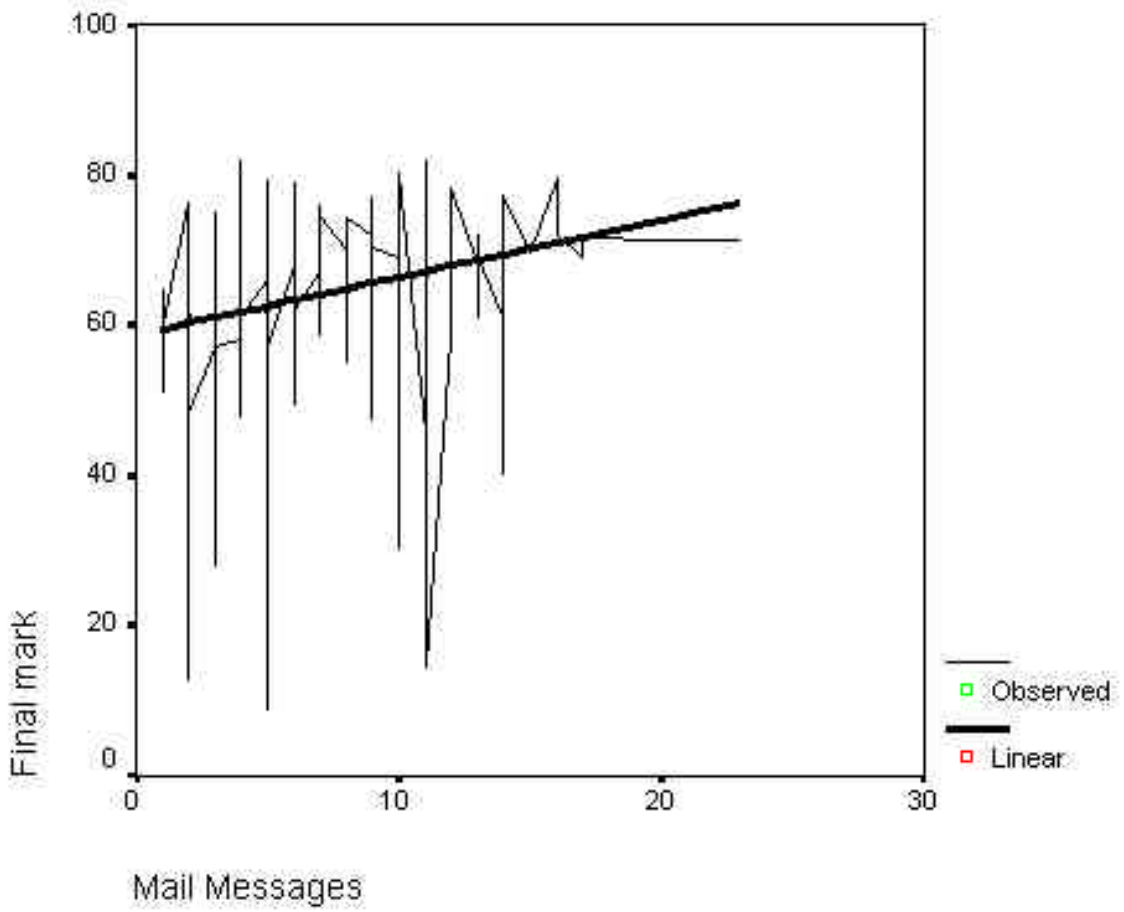


Figure 5. The linear regression line for the number of messages sent (*independent variable) and the student learning outcome in the course as measured by the final* mark (*dependent variable).*

Table 3. Table of regression coefficients for each of the BIL years

Year	Statistic			Unstandardized Coefficients		95% Confidence Interval for B	
	N	F	Sig.	B	Std. Error	Lower Bound	Upper Bound
1999	253	14.98	0.0001	52.51	2.01	48.55	56.46
				1.59	0.41	0.78	2.4
2000	186	13.06	0.0004	59.07	1.23	56.64	61.5
				0.72	0.2	0.33	1.11
2001	213	12.7	0.0005	58.71	1.59	55.58	61.85
				0.76	0.21	0.34	1.19

CONCLUSION

The BIL case study over three years shows a consistent picture of message behavior and learning outcomes. The positive relationship between messaging frequency, message length and summative learning outcomes provides confidence in the effectiveness of the teaching strategies involved. Student messaging behavior made visible in the e-mail record can be used as a formative indicator of time on task and learning quality. Such computer-automated measures can provide guidance for instructors on student progress.

REFERENCES

Geer, R. & Barnes, A. (2001, in press) *Towards a framework of teaching and learning through technology-mediated interaction*. Seventh IFIP World Conference on Computers in Education. Aarhus: Kluwer Academic.

Gunawardena, C.N., Lowe, C.A., & Anderson, T. (1997) Analysis of a global online debate and the development of interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*. 17(4), 397-431.

Henry, F. (1991) Computer conferencing and content analysis. In A.R. Kaye (ed) *Collaborative Learning through Computer Conferencing*. Heidelberg: Springer-Verlag

Hiltz, S. R. (1998) Collaborative learning in Asynchronous Learning Networks: Building learning communities. [Online] http://eies.njit.edu/~hiltz/collaborative_learning_in_asynch.htm [1999, August 8]

Johnson, D. W. & Johnson R. T. (1996) Cooperation and the use of technology. In D. Jonassen (ed) *Handbook of Research for Educational Communication and Technology*, 1017-1044. New York: Macmillan Library Reference.

King, J. L. & Doerfert, D. L. (1996) *Interaction in the distance education setting*. [Online] http://www.ssu.missouri.edu/ssu/Aged/NAERM/s-e-4.htm [1999, September 7]

Kirby, E. (1999) *'Building interaction in online and distance learning'*, Unavailable URL, see sourced at ekirby@westga.edu

Newman, D., Johnson, C., Webb, B. & Cochrane, C. (1997) Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information Science* 48(6) 484-495.

Ramsden, P. (1992) learning to Teach in Higher Education. London: Routledge.

Sringam, C. & Geer, R. (2000) An Investigation of an instrument for analysis of student-led electronic discussions In *Learning to Choose*, ASCILITE 2000 Conference Proceedings, 81-91Coffs Harbour, NSW, Dec 2000.

Schooling and intelligence: Effects of track of study on level and profile of cognitive abilities

Jan-Eric Gustafsson
Goteborg University *jan-eric.gustafsson@ped.gu.se*

Paper presented at the Spearman Conference, Sydney, 2001

The main purpose of the paper is to investigate effects of schooling on intelligence, with a particular emphasis on two questions: (1) is there a causal effect of schooling on intelligence, and if so how strong is the effect, and (2) what aspects of intelligence are being influenced? Reviews of the empirical research on the effect of schooling on intelligence are presented, and it is concluded that the research indicates that there is an effect of around 2 IQ points per year of additional schooling. Some interpret the effect of schooling on intelligence as an improvement in test performance only, there being no change in intelligence, while others view the change as a real change in intelligence. In order to obtain a better understanding of the nature of the change in intelligence as a function of schooling it is proposed that research should leave the notion of an undifferentiated concept of intelligence in favor of a multidimensional conception of intelligence. It is also proposed that rather than studying effects of an undifferentiated amount of schooling effects of different types of curricula on different aspects of intelligence should be investigated. An empirical study is presented in which changes in intelligence during two years of study on different tracks in upper secondary education are investigated for a group of some 14 000 males. At enlistment to military service at age 18 a cognitive test battery measuring Fluid ability (Gf), Crystallized intelligence (Gc) and General visualization (Gv) ability was administered. A latent variable model with five factors was fitted to the grades in the leaving certificate from grade 9 of compulsory school, and the five factors were used to control for entry differences to the tracks of upper secondary school. Effects corresponding to an improvement of about 3 IQ points are estimated for academic tracks. The academic tracks with technical and science orientation cause at least as strong an improvement in Gv, while for Gc weaker positive effects were obtained for the academic tracks and some of the vocational tracks. It is concluded that certain schooling experiences may cause improvements both in general cognitive ability, and in specific abilities.

INTRODUCTION

The field of research on intelligence is rich in paradoxes. There is, thus, ample evidence of stability of individual differences in intelligence over long periods of time, extending even from childhood to old age(e. g., Deary et al., 2000).Assumptions of stability are also made in mnay practical applications where it has proven useful to treat results obtained on tests of intelligence as representing fairly fixed characteristics of individuals that may be used for purposes of prediction, selection and diagnosis. But there also is evidence that intelligence can change, and that these changes may be quite dramatic. The best example of this is the so called secular change of intelligence, according to which there is a strong increase in the intelligence level of the population over time. This has been observed to take place since at least the 1930s, and since then the level of intelligence has risen by about one standard deviation (Flynn, 1984, 1987, 1999; Neisser, 1998). While no one challenges the empirical finding of the secular trend

there is little agreement as to the explanation and meaning of the finding (Neisser, 1998). One of the many explanations that have been put forward is that the secular trend is due to the increased general level of education, and there is evidence that attending education may affect individual level of intelligence in a positive way (e. g., Husén & Tuijnman, 1991).

The possibility that education may affect intelligence is, however, an idea that has generated much controversy. The discussion following the publication of Jensen's (1969) paper on how much intelligence may be boosted through preschool participation is an example of this. Reviews of research by, among others, Ceci (1991), Herrnstein and Murray (1994), and Winship and Korenman (1997) also arrive at quite different conclusions about to which degree schooling affects intelligence. One of the reasons for the divergence in opinions among reseachers is that the problem of determining the degree of effect of quality and quantity of schooling offers great methodological challenges.

Let me illustrate this by looking at the design and results of one of the classical studies in this field, namely the study by Härnqvist (1968a, 1968b). In this longitudinal study Harnqvist tested a representative 10% sample of the Swedish population of 13-year olds with a test battery comprising a verbal, a spatial and an inductive test. At the age of 18, the male subset of the sample took another test battery of a similar composition at the enlistment to military services. Harnqvist also had access to information about the educational careers of the subjects, which were differentiated into different tracks both in secondary school (grades 7-9) and in upper secondary school (grades 10-12; subjects still in school were tested either during the second semester of grade 11 or during the first semester of grade 12). Some subjects left school for work after grade 7, and quite a few did so after grade 9. In the analysis of the data Härnqvist used the test results at age 13 to partial out differences in entry characteristics to the different tracks of education, in order to determine treatment effects on intelligence of the differing kinds and amounts of schooling. In order to obtain more correct estimates of treatment effects in the covariance analytic design, Härnqvist devised a method for correcting the intelligence measures at age 13 for unreliability. A main finding was that students who had the most academically oriented education gained approximately .6 standard deviation units in intelligence as compared to those with the least amount of academic education. This is a fairly strong treatment effect, which exceeds the level of 0.5 stipulated for a medium effect size. Estimated on a yearly basis the effect is approximately 0.12 sd units per year, or 1.9 IQ points.

This empirical result raises, however, a number of questions concerning validity and interpretation. The most fundamental question is, of course, the one about internal validity: is the conclusion about a causal effect correct, or may this conclusion be challenged? Since this is a non-experimental study in which subjects were not randomly assigned to treatments there is the possibility that selection effects are an alternative explanation of the findings. This alternative explanation was stressed by Brody (1992), who said:

> Individuals who chose or were assigned to an academic track might have gained in IQ even if they had been randomly assigned to a less rigorous academic education. Consider two individuals with the same IQ who elect to enter different educational tracks at the secondary school level. The student who chooses the academic track may like to read books more than the individual who chooses a less academic track. Differences in intellectual interests may be related to changes in IQ.

This objection certainly is valid in principle, as it raises all the fundamental problems which are related to making causal inferences from non-experimental data. As was observed by Lord (1963) there is in such situations no dependable method for determining what is the appropriate adjustment of initial differences between groups. The only way out of this situation seems to be the recommendation made by Campbell (1963) to: "...vigorously attend to the specific plausible rival hypotheses appropriate to each situation" (p. 214). I will return to the evidence available concerning the validity of the alternative hypothesis put forward by Brody (1992).

If the Härnqvist finding is, temporarily at least, accepted to be internally valid, there is still the question of interpretation of what the finding means. Ceci (1999) argued, in essence, that there is a change of 1.8 IQ points for each additional year of academic schooling (which estimate, incidentally, agrees quite well with the estimate of 1.9 IQ points obtained by Härnqvist), but he also emphasized that the change is a change in intelligence test scores and not in intelligence: "... while schooling seems to prop up IQ test performance, it does not seem as obvious that it increases intellectual development, particularly if we conceptualise the latter in terms of novel problem solving" (Ceci, 1999, p. 171). This interpretation may be a reasonable one, but Ceci does not supply any empirical support for it. The question what the change in observed test performance signifies thus seems to be an open one.

The main purpose of the present chapter is to investigate effects of schooling on intelligence, with a particular focus on the two questions identified in relation to the Harnqvist example: (1) if it is possible to claim that there is a causal effect of schooling on intelligence, and in that case how strong the effect is, and (2) what aspects of intelligence are being influenced?

REVIEWS OF STUDIES OF EFFECTS OF SCHOOLING ON INTELLIGENCE

As has already been mentioned, several reviews of the literature on effects of schooling on intelligence have been published, and since the literature is very large, I will first review the reviews.

One of the fundamental methodological problems in research which aims to determine effects of schooling on intelligence is that strong designs which involve randomization of subjects are out of the question for practical and ethical reasons. This makes it difficult to establish casual relations, but it has spurred researchers into being creative to find other ways to establish the effects of schooling on intelligence. Ceci (1991) identified some 200 studies with eight different types of designs, and even though it is not possible to discuss all these here, it may be worthwhile to bring up some of the main designs and findings.

One class of evidence concerning the effect of schooling on intelligence is that during the summer vacation there is a small decrement in IQ, which effect seems to be stronger for children of low socio-economic status (e. g., Jencks et al., 1972). Another, similar, type of studies has investigated effects of intermittent school attendance on intelligence. In this group, Ceci (1991) refers to several early studies of children who were prevented from attending school regularly because of lack of travelling, handicap, or living in areas without access to education. These studies demonstrate a successive drop in IQ as a function of age for these groups of children. Ceci also refers to other studies showing cumulative deficits in IQ as a function of how much of school is missed.

Another class of studies investigates effects on intelligence of delayed onset of school. An example is a Dutch study (DeGroot, 1951) of the effects of schools being closed during World War II. This study estimated that when school onset was delayed several years, IQ dropped approximately 7 points. Ceci concluded that studies of effects of missed or delayed schooling show that these factors cause consistently large negative effects on IQ.

Ceci also reviewed studies investigating length of academic schooling. One example of such a study is the Härnqvist study which has already been described. There also are other such studies, several of which are Scandinavian. One is the Husen (1951; se also Husén & Tuijnman, 1991) study, in which 613 Swedish men were tested at the age of 10 in 1938, and then again at the induction to military service 10 years later. This study showed that there is an effect of completing secondary school (junior high school) and an even stronger effect of completing upper secondary school (high school). Lund and Thrane (1983) gave a large sample (N = 7 703) of Norwegian school children in grade 7 a set of military enlistment tests. Five years later the boys got the same tests at the military enlistment, leaving a sample of 2 485 individuals who had test results at both testing occasions. Lund and Thrane used several

different methods to analyze the data, and among them were similar analys of covariance techniques as were used by Härnqvist. With correction for attenuation in the pretest scores they arrived at estimates of increases of from 2.5 to 2.8 IQ points per year of education. Lorge (1945) in an early study also used a longitudinal design to study effects of amount of schooling on intelligence, and concluded that there were substantial school-related increases in intelligence.

Yet another category of studies reviewed by Ceci (1991) argue that there are effects of schooling on intelligence, because the pattern of relations between schooling variables and achievement test scores are similar to the pattern of relations between schooling variables and intelligence scores. This suggests that IQ scores are just as influenced by schooling as is academic achievement.

Ceci (1991) argued that one of the strongest designs for investigating effects of schooling on intelligence is a cohort-sequential design in which children of the same age enter school at different times (i. e., they go in different grades), and stay there for similar lengths of time. One such a study was reported by Cahan and Cohen (1989) who investigated the slope of within-grade regressions of intelligence test scores on chronological age, comparing the regressions across grades to determine the impact of years in school. They found that the effect of schooling was substantial, and, indeed, about twice as strong as the effect of chronological age.

Ceci (1991) observed that for each of the different categories of evidence reported, there is evidence that schooling affects intelligence. He also noted that it would be possible to challenge this as a causal conclusion in virtually every study, because they are all basically correlational. However:

> ... despite the many interpretive snarls one confronts with correlational data, when one considers the entire corpus of correlations that have been reported this century; the high correlations between IQ and schooling are difficult to account for on the basis of genetic selection or any other explanation (e. g., motivational differences or parental SES), because these mechanisms appear farfetched in many of the studies that were reviewed. The most parsimonious account of the correlations that have been reviwed is that of a direct causal link, namely, that the processes associated with schooling influence performance on IQ tests through a combination of direct instruction and indirect inculcation of modes of cognizing ... (Ceci, 1991, p. 711).

On the basis of a review of the literature and an empirical study, Herrnstein and Murray (1994, Chapter 17) arrived at quite a different conclusion, namely that there is little evidence that differences in the amount of schooling accounts for much of the intellectual variation. They did, though, observe that a large scale systematic attempt to raise the intelligence in Venezuela and long-term coaching on tests like the Scholastoc Aptitude Test can have an effect of the order of 3 IQ points. Herrnstein and Murray (1994, Appendix 3) also reported an empirical study of their own, in which they analyzed changes in the Armed Forces Qualification Test (AFQT; a subset of the Armed Services Vocational Aptitude Battery) for a subset of 1408 cases in a longitudinal study (NLSY), for which results were available on one or more previously administered IQ tests. The results showed an increase of about 1 IQ point per year of additional schooling. This is a positive effect, but the estimate is lower than what has been obtained in other similar studies.

Winship and Korenman (1997) reanalyzed the Herrnstein and Murray (1994) data, correcting technical problems with the data (e. g., missing data problems, and the treatment of individuals entered multiple times) and using alternative model specifications (e g., comparing models with different assumptions about the reliability of the independent variables). The reanalysis resulted in higher estimates of the effects of schooling on intelligence: according to the model preferrred by Winship and Korenman (1997, p. 231) there was an estimated effect of 2.7 points of IQ per year of education.

Winship and Korenman (1997) also reviewed the research on effects of schooling on intelligence, with a special emphasis on studies using an analysis of covariance design. They tried to obtain a quantitative estimate of the magnitude of effects of schooling from each study. These estimates covered a quite large interval, the lowest being around 1 IQ point per year of schooling and the highest being around 4 IQ points per year of schooling.

If more emphasis is placed on the Winship and Korenman (1997) reanalysis of the Herrnstein and Murray (1994) study than on the original analysis, these reviews do seem to agree that there are positive effects of at least a couple of IQ points for each year of additional schooling.

Ceci (1991) brought up possible mechansisms that may account for the effects of schooling on intelligence, but he also observed that the empirical results are often too coarse to allow a more detailed analysis. Different types of schooling practices may thus be expected to influence some types of IQ tests more than others, but studies have typically not investigated effects on tests measuring different aspects of intelligence.

Ceci (1991) also pitted two possible explanations against one another. According to one of these schooling influences intelligence because the experience of being in school alters the individuals' cognitive processes in a fundamental manner, such as through fostering more abstract and disembedded ways of thinking about the world. According to the other, schooling only supplies the individuals with knowledge relevant for responding to IQ tests, and efficient ways of responding to items in such tests. Evaluating the support for these hypotheses, Ceci seems to favour the latter one. He emphasized that schooling involves teaching answers to questions on IQ tests, and that it is in school that one is most likely to come across information relevant to IQ tests. Ceci also observed that schools teach modes of thinking that are rewarded on intelligence tests, such as using paradigmatic classifications. As a third explanation, Ceci emphasized that schools prepare students with values that support efficient test-taking.

In the absence of more specific information about how different types of schooling affect different aspects of intelligence, it does seem, however, that conclusions about the mechanisms involved are premature.

There is also another reason why it does seem necessary to focus more on effects of different types of education than has been done in previous research. This is because education nowadays cover a much greater part of life for most individuals, at the same time as the concept of education, particularly at upper secondary and tertiary levels has been broadened to include much more of vocational education.

Commenting upon a finding by Flynn (1987) that only 5 % of the Dutch population increase in intelligence between 1952 and 1972 was due to increases in schooling, Ceci (1991) observed that " ... perhaps dropping out of high school and going into work-study apprenticeships during the early half of this century has been replaced by completing high school on a nonacademic track (i. e., taking relatively little academic course work). There has been an enormous growth of occupational and alternative education during the 1960s and 1970s, and this may have resulted in many students attending school even though little academic work was accomplished, whereas in former times these same students would not have remained in school.

From the literature reviwed so far it thus seems that we may draw the general conclusion that schooling has an effect on intelligence. However, several concerns which need further attention have also been identified. There is, thus, a need for better control of factors which are correlated both with the independent variable and the residuals of the dependent variable, and which therefore are threats to the validity of the inference that it is schooling that causes improvements in intelligence. One example of such an omitted variable may be differences in intellectual interests, which was proposed by Brody (1992) as a possible alternative explanatory factor in the Härnqvist (1968) study. Another factor which needs closer attention is type of education. Ceci (1991) suggested that only academically oriented types of programs

cause effects on intelligence, and it certainly is a worthwhile question to ask to what extent there are differences in the effects of different types of programs. A third aspect which needs refinement is the conceptualization and measurement of the construct of intelligence. The large majority of studies have relied on global and undifferentiated IQ measures, but considerably more information may be gained if instead a model of intelligence is adopted which allows for multidimensionality, such as a hierarchical model (Carroll, 1993; Gustafsson, 1984, 1988, in press-a). Particularly when effects of different types of educational programs are studied it would seem necessary to adopt a more multifacetted conception of intelligence.

There is, indeed, a largely unnoticed study of effects of schooling on intelligence which improves the design in all these three respects. This is a Swedish study by Balke-Aurell (1982), which is described in somewhat greater detail below.

The Balke-Aurell Study

The study is a follow-up and extension of the Härnqvist (1968) study, and it partly relies on data from the same large-scale longitudinal study. While the main purpose of the Härnqvist study was to investigate changes in the general factor of intelligence, the main purpose of the Balke-Aurell study was to investigate changes in specific ability factors (primarily verbal/ reasoning and spatial/technical abiliies). In order to study the influence of different types of educations and occupations on the pattern of abilities, Balke-Aurell classified the educations and occupations into verbal and spatial/technical domains, and used these classifications as independent variables.

The study comprised two cohorts of data, one representative 10 per cent sample of males from the Swedish population born in 1948, and one born in 1953. The reason why the study is restricted to males is that it uses information about test results from the military enrollment at age 18. The number of cases from the 1948 cohort with complete data was 4 443 (82.6% of the originally drawn sample). For the 1953 cohort there were 3 847 cases with complete data (80.8% of the originally drawn sample).

The first wave of data collection took place when subjects were 13 years old, and typically attended grade 6. At this time a test battery with three tests (a number series test measuring inductive ability, a vocabulary test measuring verbal ability, and a spatial visualization test measuring spatial ability) was administered along with a questionnaire, asking, among other things, about preferred and actual leisure-time activities, attitudes toward school, and future plans for education and occupation.

Another wave of data collection took place when the subjects were tested at age 18 at the enlistment to military service. At this time a test battery with four tests was administered (an instructions test measuring verbal ability, a concepts test measuring verbal reasoning, a form-board test measuring spatial ability, and a mechanical comprehension test measuring spatial-technical ability). In order to form maximally comparable composites of scores from the two batteries Härnqvist (1968) used canonical correlation analysis to form one g-factor composite (g) from each battery, along with a bipolar factor contrasting verbal/reasoning ability with spatial/technical ability (v-s). Balke-Aurell used the same procedure, but she also treated the test batteries as measuring two correlated factors, one verbal (V) and one technical-spatial (T/ S).

Information about educational and occupational experience was also collected at age 18. Balke-Aurell categorized type of educational experience into three categories (educational structure): lines of study dominated by verbal subjects, those dominated by spatial/technical subjects, and a residual group not dominated by any of these. Educational experience also was classified into four educational levels: compulsory school only, practical-vocational education, lower secondary education, and upper secondary education.

Occupational experience was classified into five categories. One category comprised occupations with mostly verbal tasks (V). There were three categories of occupations making different kinds of demands of a spatial-technical nature: occupations involving spatially oriented tasks (T1); occupations involving tasks with demands on comprehension of technical-mechanical matters (T2), and occupations with tasks with demands on independent decisions concerning technical-mechanical problems (T3). The fifth occupational category did not pose any of these demands.

Balke-Aurell analysed the data by regression analysis and by fitting path models with latent variables. One analysis replicated the Härnqvist analysis on data from the 1953 cohort, with highly similar results. For the 1948 cohort the effect was .61 sd unit for the difference between those with the highest level of education as compared to those with the lowest level of education, while for the 1953 cohort the corresponding effect was .54 sd units. Balke-Aurell (1982, p. 105) suggested that the slight decrease of the effect may be due to the fact that most of the boys of the 1953 cohort attended the newly introduced comprehensive school with a less streamed curriculum than the school system attended by those born in 1948.

Changes in the v-s factor at age 18 for the different educational and occupational factors were also estimated, using the v-s factor at age 13 as control variable. These analyses revealed that for occupations at all educational levels there were quite consistent changes in the expected directions for both cohorts. The V group changed towards the verbal end of the factor, while T1, T2, and T3, generally in that order, changed towards the spatial end of the bipolar factor. The overall effect, expressed in sd units was .34 for the 1948 cohort, and .42 for the 1953 cohort. There was, however, a tendency for the effect to be stronger at higher educational levels.

For those in education until age 18 there also were changes in the expected directions for the three groups. The overall effect was .38 sd units for the 1948 cohort and .26 for the 1953 cohort. Balke-Aurell cautioned that it is difficult to compare the two cohorts because of changes of the Swedish educational system during these 5 years which makes it difficult to fit the educations into the same coding system.

Balke-Aurell also conducted multiple regression analyses in which a wide range of variables from the first wave of measurement was included as additional control of self-selection effects. Among these were school marks, social class, municipality, interests and ambitions, and social relations. Inclusion of these variables caused considerable increases in the amount of explained variance, but the estimates of treatment effects generally were unaffected. For the students the effect for the 1948 cohort increased slightly, from .38 to .39, while for workers the effect dropped from .34 to .28.

It was noted, however, that several of the variables introduced into the multiple regression analysis had low reliability, which causes bias in the estimates. The correction introduced and used by Härnqvist (1968a) could not be applied in this situation. This was one reason why Balke-Aurell also performed a series of analyses using latent variable models. Another reason for fitting such models was that they allow specification of two correlated latent ability variables at age 18, which would make it possible to study not only whether there has been a relative change in the balance of a bipolar ability, but also which ability has been affected.

These models were path models in which Verbal-scholastic ability, Verbal interest, Spatial ability, Technical interest and Social background were independent variables, Educational level and Educational structure were intervening variables, and Verbal ability and Technical/ spatial ability at age 18 dependent variables. According to this model there was no effect of Educational level or of Educational structure on the Verbal ability factor. However, for the Technical/spatial factor there was a significant standardized partial regression of .07 for Educational level, and a coefficient of .08 for Educational structure.

A models also was fitted in which the g and v-s factor were used as dependent variables. This model showed the g-factor to be affected by Educational level (standardized regression coefficient .07), and the v-s factor to be affected by Educational structure (.11).

Balke-Aurell concluded that the higher the educational level, the stronger is the increase in general intelligence. She also concluded that specific ability factors develop in accordance with verbal and technical types of education, and, to a lesser extent, with type of occupation. These effects were mainly found, however, for spatial/technical ability, and to a lesser extent for verbal ability.

The Present Study

The Balke-Aurell study demonstrates that it is necessary to take into account the fact that education is not a homogeneous activity, but that the characteristics of different educational tracks are important for effects on intelligence. The study also shows that it is necessary to investigate differential effects of schooling on different aspects of intelligence. Below an empirical study will be presented which aims to extend the study of effects of different educational programs, on different aspects of intelligence.

Sweden has 9 years of compulsory education, and 3 years of voluntary upper secondary education (or high school). However, even though upper secondary education is not compulsory, about 90 % of the cohort participates in 12 years of education (OECD, 2001). This increase in participation rate compared to the cohorts studied by Härnqvist (1968a, 1968b) and Balke-Aurell (1982) is mainly due to an expansion of vocational education at the upper secondary level. This had started for the 1954 cohort, who went through upper secondary school in the early 1970s, but the great expansion occurred after that. Until the mid 1990s most vocational programs comprised two years of study, but after this time all upper secondary education programs in Sweden comprise three years of study, and have a higher level of ambition as to academic content.

The current study investigates the cohort born in 1976. Normally students born this year left compulsory school at the age of 16 (school starting age is 7 in Sweden) in 1992 and started upper secondary school in the autumn the same year. At this time the old system of 2-year vocational programs was still in operation, but in some parts of the country the new 3-year programs had already been implemented. This makes it possible to make some comparisons between the vocational programs of different types.

This study also takes advantage of the fact that a new military enlistment battery was introduced in 1994. During the latter part of this year the old test battery consisting of four paper and pencil tests, which were designed to measure the general factor only, was replaced by a computer administered test, consisting of 10 subtests designed to measure the factors Fluid intelligence (Gf), Crystallized intelligence (Gc) and General visualization (Gv) (Carlstedt 2000; Mårdberg & Carlstedt, 1998). The three factors are estimated within a hierarchical modeling framework, according to which the Gf-factor is identical with the g-factor (Gustafsson, 1984, 1988, in press-a) and Gc and Gv correspond to second-stratum factors of Carrolls (1993) model.

This new test battery provides much more information about different aspects of intelligence than the battery studied by Balke-Aurell. What is of special interest to investigate in the current study is, of course, if schooling affects the g-factor, or if it affects the factors at lower levels of the hierarchical model, as was hypothesized by Ceci (1999). It is also of great interest to investigate if programs of education with special curricular emphasis affect special abilities, in a continuation of the investigation conducted by Balke-Aurell (1992).

The empirical study relies on information from official registers only, while the Härnqvist and Balke-Aurell studies used data collected through tests and questionnaires as well. Register information can easily be obtained for the entire population, so this method of data collection

brings the advantage that a large set of data is available for analysis. The disadvantage is that certain types of information are not available, such as test scores and information about interests. However, there is information about school achievement and socio-ecenomic status, so a rather high degree of control of intial differences may be achieved.

METHOD

The design of the empirical study is fairly simple. Differences in results on the military enlistment test battery at age 18 for students who have followed different tracks in upper secondary school are analyzed, controlling for initial differences in grades and socio-economic background. To the extent that the control for differences due to self-selection is successful this design makes it possible to make causal inferences about effects of track of study on level and profile of intelligence.

Subjects

The data from the present study are taken from a large data base consisting of everone born between 1972 and 1979 living in Sweden in 1996. The data base includes a large number of variables taken from different registers containing information about educational choice and educational achievement (see Gustafsson, Andersson & Hansen, 2000, for a description in Swedish). Because every person in Sweden has a unique identification number which is used in almost all registers it is possible to create longitudinal data bases by combining the register data at individual level. Most of the registers are kept by Statistics Sweden, where the combination of the different registers has been conducted, after which the data has been delivered to Göteborg University in anonymous form.

The data to be analyzed here are from the 1976 cohort. This is the only cohort for which information is available both about track of study at upper secondary school and results on the new version of the military enlistment test battery. This cohort also is interesting to study because it represents a large variety of upper secondary programs. As has already been mentioned a new version of the upper secondary school curriculum was successively implemented until the mid 1990s. In the 1996 cohort there thus are persons who have followed the older version of the upper secondary school curriculum, in which vocational lines comprise two years of study, and there are persons who have followed the new curriculum, in which vocational programs comprise three years of study, and have a higher level of amibition when it comes to academic subject matters.

Because enlistment to military service in Sweden is compulsory for males only, the analyses reported here will be restricted to this gender. This not only limits the generalizability of the findings to males only, but it also entails a restriction on which programs of study can be investigated, because in many instances choice of program is highly correlated with gender. This is an unfortunate but inescapable consequence of relying on data collected for other purposes than research.

The data base comprises 52 113 males born in 1976, while the analyses to be reported are based upon a subset of 13 903 cases. The main reason why so many cases are lost is that only 17 588 cases have a score on the 10 subtests of the test battery analyzed here, which is because the new enlistment test was operational only from August 1994 (Mårdberg & Carlstedt, 1998). Only few cases were not tested at all. Those who are not Swedish citizens are excluded from the military enlistment procedure, as are those who have been diagnosed to be mentally retarded. With these restrictions, there is little reason to believe that those taking the new enlistment test do not form a representative sample from the population.

Another reason why not all cases are analyzed is that some attend programs of study which are too small to be analyzed. Because of the successive implementation of the new curriculum for the upper secondary school there were during the years 1992 to 1994 an unusually large

variety of upper secondary school tracks. Not only did the old and new curricula co-exist in the country, as a function of whether the local authorities had decided to adopt the new curriculum or not, but in some school districts an experimental pre-cursor to the new curriculum, with 3-year vocational programs, was in operation. These experimental programs are not so well defined as the others, so they have been left out of the analysis.

Information about the socio-economic background of the students is available from the 1990 Census. Parental occupation has been used to classify the background of the students according to the so called Socio-Economic Index, which here involves a categorization into one of eight groups (e. g., unskilled labor, positions requiring academic education). In the analyses group belongingness has been represented as a vector of dummy variables.

Treatments

The old version of the upper secondary school curriculum consisted of five academic lines (Liberal Arts, Economics, Social Science, Natural Science and Technology), which prepared for entry to university level education. These programs all comprised three years of study. There also was a rather large number of vocational programs, which all comprised two years of study, and which did not allow entry to university level education without further education at the upper secondary level. The new system consists of 17 national programs of study and a large number of local programs. These programs are all three years long, and they all provide formal access to university level education, even though the range of available educational programs varies highly. When the students born in 1976 entered upper secondary school the new system had been decided upon, and in some parts of the country implementation started before the official starting date (i. e., the academic year 1994/95). A smaller part of the cohort thus have their upper secondary education from one of the new programs. For most of the programs the number of students is, however, to small to allow analysis.

Table 1 presents the number of students born in 1976 who have graduated from different lines of the old upper secondary school curriculum, and for students who have graduated from some of the programs of the new upper secondary school curriculum. For the analyses only lines and programs have been included for which at least 100 students is available for analysis. This implies, regrettably, that the Liberal arts program, for which only 90 cases are available is excluded. However, when attempts were made to include this program in the models, they failed to converge, so it is necessary to exclude the Liberal arts program from the analyses.

For most programs the proportion of cases who have data on the variables to be analyzed is around 30 per cent. For students from the new programs the proportions tend to be somewhat higher. One explanation for this may be that the new programs were implemented earlier in urban areas, where also the proportion of students who have complete data for grades as well as the SEB is larger than in other areas of the country. About 25 per cent of the students who miss information about line of study have grades from compulsory school and results on the SEB. The reason why information is missing is that they have either not entered upper secondary school, or have entered but not completed their education at the age of 20. The reason why there is a larger attrition from this group than the others is that it includes a larger proportion of cases with a non-Swedish background, who are less likely to take the SEB, and who are more likely to miss one or more of the grades from compulsory school.

In the analyses to be reported here all the different lines and programs will be analyzed as different treatments, and comparisons will be made with the group of cases for which no information is available about their upper secondary education. To understand the nature of possible treatment effects it would be desirable to have a fairly detailed description of the amount of time spent on different subject matters. Given the large number of tracks this is not possible, but it may be useful to have a more detailed presentation of one example of each of the main categories of educations.

Table 1. Number of male students born 1976 from different lines and programs of study

	Graduated	Included	Percent
Two year lines			
Building and Construction Line	2346	695	29.6
Consumer Line	148		
Distribution and Clerical Line	1016	297	29.2
Electro-Telecommunications Line	2874	924	32.2
Food and Manufacturing Line	721	239	33.1
Horticultural Line	194		
Motor Engineering Line	1910	601	31.5
Music Line	181		
Nursing Line	135		
Operation and Maintenance Line	690	232	33.6
Social Line	1609	503	31.3
Processing Line	36		
Woodwork Line	343	120	35.0
Workshop Line	972	372	38.3
Three year lines			
Economics Line	3861	1136	29.4
Liberal Arts Line	311		
Natural Sciences Line	3458	1007	29.1
Social Sciences Line	3717	1101	29.6
Technology Line	5356	1551	29.0
Three year programs			
The Construction Programme	495	215	43.4
The Electrical Engineering Programme	1191	445	37.4
The Vehicle Engineering Programme	834	271	32.5
The Business and Administration Programme	961	309	32.2
The Industrial Programme	822	324	39.4
The Use of Natural Resources Programme	843	280	33.2
The Natural Science Programme	1425	488	34.2
The Social Science Programme	1913	634	33.1
No information	8404	2087	24.8

The Electro-Telecommunications line is the largest among the two-year lines. This line had during the two years 160 lessons (40 minutes) of Swedish and 240 lessons of an optional theoretical subject. For electro- telecommunications engineering 2360 lessons were alloted, divided between work technique and vocational theory, with a stronger emphasis on the former than on the latter. This curriculum thus has a weak emphasis on general academic subjects, and a strong emphasis on practical skills.

This curriculum may be compared with the new Electro- Telecommunications program. In the new upper secondary school the curriculum is based on courses, which are of three different kinds: core courses, which are taken by everyone, program specific courses, and individual choice courses. Each course has a number of credit points assigned to it, and currently a total of some 2 500 points is required for the three years. The core courses, almost all of which are in different academic subjects, comprise 750 points, and the program specific courses about the same number of points. Some 900 points may be taken as individually elective courses, which may be either general academic courses (e. g., mathematics, English) or program related courses (e. g., electronics, computer technology). In addition a project work must be completed. The proportion of academic courses thus is higher in the new program, even though the large number of elective courses makes it impossible to identify a single common curriculum. However, according to the recommended time table specified in 1997/98, 1370 hours (60 minutes) should be devoted to vocational subjects, while 680 hours should be

devoted to core subjects. Thus, even though there is a stronger emphasis on academic subjects in this program than in the two-year program, the main emphasis still is on vocational skills.

For the three-year Natural Science line, to take yet another example, approximately half the time was allotted to mathematics, science subjects and technology, approximately 25 per cent of the time to Swedish and foreign languages, and the remaining time roughly evenly split between social subjects and other subjects. The main emphasis of this program thus is academic.

In the analysis each different track will be represented by a dummy variable of its own, using the No information group as the reference group.

Grades from compulsory school

As has already been pointed out there is a strong need to identify and control for possible self-selection of students into the different tracks. This will be done through relying on the grades in 17 different subject matters given on the leaving certificate from compulsory school. The grades were assigned on a scale from 1 to 5 according to a norm-referenced grading system, in which national tests were used to achieve comparability of grades from different schools.

Andersson (1998) has fitted a five-factor model to the grades data, which holds promise not only to capture differences in general level of school achievement for the different lines and programs of study, but also differences in the profile of performance over different areas. Andersson fitted her model to the population data for students born in 1972, and the standardized factor loadings for males are shown in Table 2, along with the results for the current sample.

Table 2. Standardized factor loadings in the five-factor model for grades

	SchAch		**Non-Verb**		**MathSci**		**Lang**		**Ad**	
	Males (72)	Current	Males (72)	Current	Males (72)	Current	Males (72)	Current	Males (72)	Current
Child studies	0.78	0.79							0.23	0.30
Art Education	0.49	0.50	0.25	0.25			0.12	0.12	0.14	0.12
English	0.73	0.73			0.10	0.12	0.42	0.41		
Domestic science	0.72	0.72	0.19	0.21					0.25	
Physical education	0.39	0.44	0.23	0.21						
Mathematics	0.75	0.75	0.14	0.16	0.29	0.28	0.12	0.14		
Music	0.63	0.62	0.10	0.12			0.19	0.19	0.13	0.12
Biology	0.85	0.84	0.08	0.11	0.24	0.27			0.11	0.11
Physics	0.83	0.82	0.15	0.17	0.34	0.36				
Chemistry	0.84	0.84	0.09	0.11	0.31	0.34			0.08	0.07
Technology	0.59	0.60	0.36	0.35	0.22	0.23				
Geography	0.88	0.89	0.03	0.02						
History	0.90	0.90							-0.09	-0.08
Religious studies	0.90	0.90							-0.02	-0.04
Civics	0.90	0.90								
Crafts	0.42	0.44	0.59	0.58					0.11	0.12
Swedish	0.83	0.84					0.25	0.24	0.04	0.07

The estimated factor loadings generally are very close. As may be seen from the structure of the estimated factor loadings, the model fitted is an orthogonal model with so called nested-factors (Gustafsson & Balke, 1993). There is a general factor (*SchAch*), which has positive and fairly strong relations to all the grades. The highest loadings are observed for social science subjects which require heavy reading and much home-work. Andersson (1998) interpreted this factor as involving a strong component of verbal ability, and also a strong motivational component.

There is also a rather broad factor which Andersson labeled NonVerb, and which is most highly related to technology, crafts, and physical education, but also to several other subjects which involve figures or numbers. Andersson (1998) interpreted this to be a spatial-practical factor. The third factor relates to grades in mathematics and in science subjects, and was labeled MathSci by Andersson (1998). The fourth factor has fairly strong relations to the grade in English and in Swedish, and weaker relations to grades in Art education, mathematics, and Music. This was interpreted as a language (Lang) factor by Andersson (1998). The fifth factor, finally, has low relations to a fairly broad range of grades, and it was labelled the aesthetic-domestic factor (Ad) by Andersson (1998).

These five factors will be used to control for possible selection and self-selection effects into the different lines and programs.

The Swedish Enlistment Battery

Until 1994 the enlistment battery consisted of four tests (one inductive, one verbal, one spatial and one technical; see Carlstedt, 2000), which where combined to measure a general factor. In August 1994 this batttery was replaced by a computer administered test, consisting of 10 (or rather 12, if two experimental tests are counted as well) subtests designed to measure the factors g, Gc and Gv within a hierarchical modeling framework (Carlstedt 2000; Mårdberg & Carlstedt, 1998). This test is referred to as the Swedish Enlistment Battery (SEB). The following tests are included:

- *Synonyms 1* and *Synonyms 2*, with 25 and 20 items, respectively. These are multiple-choice tests with four or five options, from which the synonym of a given word is to be selected. There is also a test called *Opposites* (25 items) in which the task is to select the antonym of a given word. These three tests have been shown to measure Crystallized Intelligence (Gc).
- *Figure Series* (20 items) presents sequences of four figures, and the task is to complete the series by selecting two figures out five given ones. In the *Groups* (20 items) test items five figures are presented, and the task is to identify the figure that does not fit thematically. These two tests have been shown to measure Fluid Intelligence (Gf).
- In the items in *Dice 1* and *Dice 2* (20 items each) two cubes, on which three surfaces are visible, are presented. According to the instructions there is a unique symbol on each side of the cube, and on identical cubes, the symbols are placed in the same relation to each other. The task is to find out if the two cubes, if turned, could be identical, or if they are different. In the *Metal folding* test (16 items) a drawing of an unfolded piece of metal is presented and the task is to find the three-dimensional object out of four which corresponds to the two-dimensional drawing. The items in the *Block rotation* test (20 items) present a three-dimensional target object, and the task is to select the identical three-dimensional rotated object out of five. In *Technical comprehension* (16 items) the items all constitute illustrated technical and physical problems, and one out of three suggested solutions should be selected as the correct one. These five tests have been shown to measure General visualization (Gv). The Technical comprehension test has a loading on Gc as well, and there also is overlap between the specific components of Dice 1 and Dice 2 (Mårdberg & Carlstedt, 1998).

Factor models have been fitted to this test battery in several previous studies (e. g., Carlstedt, in press; Carlstedt, Gustafsson, Ullstadius, 2000; Mårdberg & Carlstedt, 1998). These models have genereally been taken to be hierarchical models of the nested-factor kind (Gustafsson, in press-a; Gustafsson & Balke, 1993) in which a general factor factor (g) is related to every test and Gc' and Gv' are introduced as residual factors to account for the remaining covariance between these two groups of tests. The model does not include a residual Gf-factor, because of the empirical equivalence between g and Gf (Carroll, in press; Gustafsson, 1994, in press-a).

Such a model has been fitted to these data as well, and the standardized factor loadings are shown in Table 3.

Table 3. Standardized loadings of the tests in the SEB on the three factors of the NF-model

	g	Gv	Gc
Figure series	0.89		
Groups	0.80		
Dice 1	0.70	0.20	
Dice 2	0.68	0.29	
Metal folding	0.71	0.38	
Block rotation	0.69	0.34	
Technical comprehension	0.68	0.24	0.17
Synonyms 1	0.61		0.64
Synonyms 2	0.66		0.59
Antonyms	0.68		0.59

As may be seen there are substantial loadings of all the tests on the g-factor, and particularly so for the two Gf-tests. The tests hypothesized to load on Gv' do so, but the loadings tend to be relatively low, which indicates that it is difficult to separate the g- and Gv'-factors in this battery of tests. For Gc' the three vocabulary tests have quite substantial loadings. Being restricted to vocabulary tests only it may be noted that the Gc factor is a fairly narrow factor.

The model includes a covariance between the residuals of Dice 1 and Dice 2 as well, and when this covariance is introduced the model fit is quite acceptable, the RMSEA being 0.044, with a narrow confidence interval.

This hierarchical model will be used to estimate the g-factor. However, this model is less useful for investigating efffects of schooling on Gc and Gv, because g is partialled out from these dimensions, and g may be affected by treatment effects. For purposes of analysis of effects of schooling on the abilities at lower levels of the hierarchy an oblique measurement model will be used as well. This model specifes three correlated factor Gf, Gv, and Gc, with the pattern of standardized loadings shown in Table 4.

Table 4. Standardized loadings of the tests in the SEB on the three factors of the oblique model

	Gf	Gv	Gc
Figure series	0.90		
Groups	0.80		
Dice 1		0.74	
Dice 2		0.74	
Metal folding		0.79	
Block rotation		0.77	
Technical comprehension		0.56	0.23
Synonyms 1			0.87
Synonyms 2			0.89
Antonyms			0.91

In this model there is no g-factor, and the Gf-factor is measured by Figure series and Groups alone. Otherwise the pattern of loadings matches that of the hierarchical model. In the oblique model there are, however, high correlations among the factors. Gf and Gv thus correlate 0.90, Gc and Gv correlate 0.69, and Gf and Gc correlate 0.72. This model too fits well, with an RMSEA of 0.046.

RESULTS

The analysis and reporting of results proceeds in three steps. First relations between the factors of the model for the grades and the SEB-factors will be investigated. Next relations between tracks of study and the grades will be studied, in order to investigate the extent of selection and self-selection to the tracks. In the third step relations between lines of study and the factors of the SEB-models are investigated with control for grades and socio-economic status.

Relations among grades from compulsory school and test performance at 18

Unless there is a strong relationship between the grades from compulsory school and performance on the SEB the grades cannot be used to control for entry differences in level of performance among the different tracks. It is thus necessary to investigate the amount of relationship among these two sets of factors, which has been done by regressing, in two separate analyses, the g-factor and and the factor of the oblique SEB-models onto the five factors of the model for the grades. The standardized regression coefficients are presented in Table 5.

Table 5. Standardized relations among the latent variables for the grades and SEB models

	SchAch	Non-Verb	MathSci	Lang	Ad	Expl var
g	0.64	0.19	0.31	0.30	-0.09	64%
Gf	0.63	0.19	0.33	0.28	-0.09	60%
Gv	0.52	0.36	0.36	0.24	-0.12	60%
Gc	0.71	-0.09	0.22	0.36	-0.05	70%

The pattern of results are highly similar for the g- and the Gf-factors, and no less than 64% of the variance in g is accounted for. The general SchAch factor contributes most, but MathSci and Lang also have substantial relations to g. For Gv there is a strong positive correlation with SchAch and Lang, which is because there is a very substantial amount of g-variance in this factor, but this factor also has rather strong relations with NonVerb and MathSci. For Gc there is a very strong relation of .71 to SchAch, and there also are relations to Lang and MathSci.

Between 60% and 70% of the variance in abilities is accounted for by the grade factors. In spite of the fact that intelligence scores are not available the amount of variance accounted for is at least as high as in previous studies. This is because a latent variable modeling approach is used which takes full advantage of the multidimensionality of grades and SEB tests, and which allows estimation of error-free latent variables. The problem of disattenuation of relations by taking unreliability into account is thus not present here.

Tracks and grades

In the Swedish school system grades have traditionally served an important function as a tool for selection into upper secondary school from compulsory school. During the 1990s, however, little selection has been necessary because by and large study places have been available to match the demand. However, even though little explicit selection is made on the basis of grades, processes of self-selection may cause considerable differences in means of grades over the different tracks in upper secondary school.

In order to investigate such differences the five factors of the model for the grades have been regressed into the full set of dummy variables for the lines of study, and the partial regression coefficients have been estimated. This analysis expresses on a convenient scale, which is comparable across the different dimensions, track differences in the level of the five factors. As has been observed by Lubinski and Humphreys (1996) correlations are often misinterpreted to indicate weaker relations and smaller differences than other measures, such as effect size measures. However, they also demonstrate that correlations and measures which express mean

differences can be translated into one another, so such misinterpretations can be avoided. Some simple rules of thumb are obtained from knowing that a small effect size (> .20) translates into a correlation of .10, that a medium effect size (> .50) corresponds to a correlation of .24, and that a large effect size (> .80) translates into a correlation of .37.

As may be seen in Table 6 choice of, and successful completion of, track accounts for no less than 55 % of the variance in SchAch. This is primarily due to the fact that the academically oriented three year lines have a much higher mean on the SchAch factor than have the vocationally oriented two-year lines. Among the new programs the Natural Science Program and the Social Science program also have a higher mean on the SchAch factor, but not as high as the old lines. The two-year Electro-Telecommunications line and the three-year Electrical Engineering program have a higher level on SchAch than the other vocational programs.

Table 6. Standardized coefficients for the differences in grades between lines

	SchAch	Non-Verb	MathSci	Lang	Ad
Two year lines					
Building and Construction Line	0.05	0.15	-0.01		0.04
Distribution and Clerical Line	0.04	0.01	-0.01		0.04
Electro-Telecommunications Line	0.15	0.19	0.12		0.08
Food and Manufacturing Line	0.06	0.03	0.00		0.09
Motor Engineering Line	-0.01	0.12	0.01		0.03
Operation and Maintenance Line	0.03	0.09	0.02		0.00
Social Line	0.10	0.02	-0.02		0.03
Woodwork Line	0.00	0.12	-0.01		0.02
Workshop Line	-0.07	0.04	-0.01		-0.02
Three year lines					
Economics Line	0.36	0.07	0.07		0.11
Natural Sciences Line	0.53	0.06	0.32		0.07
Social Sciences Line	0.43	0.02	0.02		0.05
Technology Line	0.50	0.21	0.32		0.16
Three year programs					
The Construction Programme	0.01	0.06	-0.03		0.03
The Electrical Engineering Programme	0.10	0.10	0.10		0.07
The Vehicle Engineering Programme	-0.02	0.08	0.03		0.01
The Business and Administration Programme	0.05	0.02	-0.03		0.03
The Industrial Programme	-0.03	0.06	0.00		-0.02
The Use of Natural Resources Programme	0.04	0.07	0.00		0.01
The Natural Science Programme	0.31	0.09	0.24		0.07
The Social Science Programme	0.25	0.02	0.05		0.05

In the MathSci factor 24 % of the variance is accounted for, and this is almost entirely due to the fact that the Natural Sciences line, the Technology line, and the Natural Sciences program have a high mean on this factor. Here too the Electro-Telecommunications line and the Electrical Engineering program have intermediate means.

The lines of study account for 8% of the variance in the NonVerb factor. The Technology line, the Electro-Telecommunications line, the Building and Construction line, the Motor Engineering line, and the Woodwork line all have correlations higher than .12.

For the Lang factor 11.6% of the variance is accounted for. The academically oriented lines and programs have a higher level on this factor. Had the Liberal Arts line been included in the analysis this line would have been seen to have a considerably higher mean on this factor.

For the Ad factor only 4% of the variance is accounted for, and there does not seem to be any clear pattern of differences among the lines with respect to this factor.

The results presented in Table 6 show that the different tracks have widely varying levels of entry performance, not only on the general school achievement factor, but also on the narrow achievement factors. It is quite interesting to see that these differences go along with the different contents and requirements of the lines. Thus, for the programs with an emphasis on mathematics and science the MathSci factor is high, and for programs in which spatial-practical skills are important the NonVerb factor is high. Unless these differences, which primarily are due to self-selection, are controlled for, they will be confounded with treatment effects.

Effects of track of study on test performance at age 18

In the final step of the analysis a model has been constructed in which the intelligence factors at age 18 have been regressed onto the dummy variables representing track and the latent variables representing individual differences in the compulsory school grades and the SES variables have been included as control variables. The model thus estimates the direct effects of line of study on test performance at age 18, controlling for entry differences.

The parameter estimates are presented in Table 7. Only significant ($p < .01$) parameter estimates are shown.

Table 7. Standardized coefficients for the differences the SEB factors between tracks, controlling for grades and ses

	g	Gf	Gv	Gc
Two year lines				
Building and Construction Line	-0.03	-0.03	-0.03	-0.03
Distribution and Clerical Line			-0.03	-0.03
Electro-Telecommunications Line	0.04	0.04	0.04	0.03
Food and Manufacturing Line				
Motor Engineering Line				
Operation and Maintenance Line	0.02	0.03		0.04
Social Line				0.02
Woodwork Line				
Workshop Line				
Three year lines				
Economics Line	0.09	0.10		
Natural Sciences Line	0.09	0.12	0.11	0.07
Social Sciences Line	0.07	0.07		0.08
Technology Line	0.14	0.17	0.17	0.10
Three year programs				
The Construction Programme				
The Electrical Engineering Programme	0.02	0.02		
The Vehicle Engineering Programme				
The Business and Administration Programme				-0.03
The Industrial Programme				
The Use of Natural Resources Programme				0.05
The Natural Science Programme	0.08	0.10	0.10	0.05
The Social Science Programme	0.06	0.07	0.03	0.03

For the g and Gf factors positive effects of attending an academic program may be observed. The strongest effect is obtained for the Technology program, but a standardized regression coefficient around .10 or slightly lower is observed for all the academic programs, including the new programs. There are also weak positive effects on g/Gf of attending the Electro-Telecommunications line and the Electrical Engineering program.

For Gv there are effects of the Natural Sciences line (0.11), the Technology line (0.17), and of the Natural Sciences program (0.10). The latter program may be regarded a combination of the Natural Sciences line and the Technology line.

For Gc there are effects of the academic lines and programs, with correlations varying between .03 and .11. There also are very weak effects, both positive and negative ones, of some of the vocational lines, but these do not seem to form an easily described pattern.

DISCUSSION AND CONCLUSIONS

The results indicate that there are effects on the factors of intelligence of schooling at the upper secondary level. However, the effects seem to be restricted to the academic programs.

The estimates for the general factor are marginally stronger when the Gf-factor in the oblique model is taken to represent this factor, than when it is estimated as a truly general factor in the NF-model. There also is some variation between the programs, the lowest effect being observed for the Social Sciences program and the strongest for the Technology program. However, the mean effect amounts to about .10, which translates into an effect size of 0.20 (Lubinski and Humphreys, 1996). Assuming an IQ scale with a standard deviation of 16 this implies a change of 3.2 IQ units during the approximately two years that the treatment lasted, or 1.6 IQ units per year. This is quite close to the estimated increase of 1.8 IQ units for each additional year of academic schooling reported by Ceci (1999), but it is at the lower end of the interval of estimates of 1 - 4 IQ points arrived at by Winship and Korenman (1997).

There is no sign that there would be any effect of the higher academic level of amibition of new 3-year vocational programs. However, it must be emphasized that this increased level of ambition has implied that the length of the programs have been extended from two to three year, while during each of the years the proportion of academic subject matter is still fairly small compared to vocational subjects. The present data was collected during the autumn semester of the third and last year of upper secondary school, which implies that for most subjects there remained almost a full year of education in the track. These data thus were collected to early to tell about the full effect of of this change.

For Gv there are fairly strong effects of the three academic programs which have an orientation towards technicology and science. For the Electro-Telecommunications line there is also a weak effect of .04, but otherwise there is no effect on this ability of the vocational tracks. Balke-Aurell (1982) found that both educations and occupations with a spatial/technical content caused a shift in the ability profile in the direction of the spatial/technical end of the bipolar v-s factor, but she also found that the effects were stronger at higher levels of education. The present results suggest that there is an effect on Gv only for the most academically oriented tracks.

For Gc there also are effects of the academic tracks, except for the Economics line, even though these effects tend to be somewhat weaker than for the other to factors of intelligence. There also are still weaker effects of some of the vocational tracks (Electro-Telecommunications line, Operation and Maintenance Line, Social line, Use of Natural Resources Program). While for Gv there seems to be a fairly strong effect for certain specific programs, the effects of Gc are weaker and associated with a wider range of programs. It must be observed, however, that the fact that the tracks under study here are those which are popular among males implies a restriction of which tracks are actually included in the analyses. The Liberal Arts program is an example of a program which too few males had selected to allow inclusion in the study. Had it been possible to include this program, a strong effect on Gc might have been found of this track. Had females been included in the study as well, a wider range of vocational tracks would also have been included in the study, some of which may have had a strong influence on Gc.

When interpreting the Gc findings it must also be borne in mind that the Gc-dimension in the present study is quite narrowly defined as a vocabulary factor. It seems, however, that it would have been difficult in the present study to adopt such a broad definition of Gc as to also include achievement in subject matter areas (Carroll, 1993), becuase then specific content taught in the different programs might have been misinterpreted as general effects on intelligence.

These results thus indicate that participation in an academic track at upper secondary school has a positive effect on Gf and, to a lesser extent, on Gc. Furthermore, academic tracks with an orientation towards technology and science affect Gv positively. This shows that even after the length of education has been extended so that the natural variation in the number of years of education up to age 18 has been eliminated, there are so large differences between the content and methods of different tracks that there are effects on intelligence which are closely similar to those obtained when variation in years of schooling were investigated.

One of the main questions in the research on effects of schooling on intelligence is how the change should be interpreted. One hypothesis is that the change is a real and profound one, which fully reflects the meaning of the construct of inytelligence. Another hypothesis is that the observed change only is a change in test performance without any concomitat change in the underlying ability. Ceci (1999) expressed the latter hypothesis as follows: "... while schooling seems to prop up IQ test performance, it does not seem as obvious that it increases intellectual development, particularly if we conceptualise the latter in terms of novel problem solving" (Ceci, 1999, p. 171). The results show quite clearly, however, that the strongest effect is observed for the Gf-factor, there being a marginally weaker effect on the g-factor which is estimated from the full set of tests in the battery. The Gf-factor is measured by tests of inductive reasoning and it clearly involves novel problem-solving, which provides support for the idea that the effects of schooling are not surface-level effects due, for example, to acquisition of specific pieces of information. One possible explanation for the positive effect on problem-solving abilities may be that the academic curricula to place an emphasis on problem solving rather than on knowledge acquisition.

The weaker effects that have been obtained for Gc should, however, probably be accounted for in terms of the increased opportunities in education to acquire knowledge about the measning of new words.

It seems, however, that the estimates of the size of effects of the current study are somewhat smaller than what has been obtained in previous Scandinavian research. One hypothesis to account for this is that the effects of schooling are diminishing with age. It seems that the highest estimates of the effects of one year of education have been obtained in studies of young students, such as the Cahan and Cohen (1989) study in which grades x to x were investigated. The present study only covers the last two years of the developmental period up to 18, while most previous study have covered a wider time span, so if the effects of education on intelligence are diminishing with age lower estimates would be expected in the present study.

It could, of course, be asked if the design of the current study allows causal inferences to be made. Given that this is a correlational study, with a less than perfect statistical control of initial differences, the results could be rejected as being expressions of mechanisms of self-selection, rather than treatment effects. With Ceci (1991) it could, however, be argued that "... in an inductive enterprise ... the usual inferential processes are relevant, namely that if studies using different methodologies, different mental measures, and different samples all converge on the same conclusion and if these studies possess uncorrelated weaknesses, then one can infer that their collective power is greater than that of their individual conclusions (Ceci, 1991, p. 718).

Compared to previous studies using the analysis of covariance design the present study employs a weaker set of measures to control for entry differences in that only school

achievement and no IQ measure is being used. However, this is partially compensated for by employing a multidimensional latent variable model which fully utilizes the information available, and which uses latent independent variables in which there is no attentuation of relationships due to errors of measurement. The amount of variance accounted for in the dependent latent variables therefore rather seems to be higher (60 % to 70%) than in most previous studies.

REFERENCES

Andersson, A. (1998).The dimensionality of the leaving certificate in the Swedish compulsory school. *Scandinavian Journal of Educational Research, 42*(1), 25-40.

Balke-Aurell, G. (1982). *Changes in ability as related to educational and occupational experience*. Goteborg, Sweden: Goteborg Studies in Educational Sciences.

Brody, N. (1992). *Intelligence (2nd ed)*. New York: Academic Press.

Carlstedt, B. *Cognitive abilities – aspects of structure, process and measurement*. Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Carlstedt, B. (in press). Differentiation of cognitive abilities as a function of level of general intelligence. A latent variable approach. *Multivariate Behavioral Research*.

Carlstedt, B., Gustafsson, J.-E, & Ullstadius, E. (2000). Item sequencing effects in the measurement of fluid intelligence. *Intelligence, 28*(2), 145-160.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components?: A reassessment of the evidence. *Developmental Psychology, 27*, 703-722.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development, 60*, 1239-1249.

Campbell, D. T. (1963). Factors relevant to the validity of experiments in social settings. In C. W. Harris (Ed.) *Problems in measuring change.* Madison: Wisconsin University Press.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. *Intelligence, 28*. 49-55.

DeGroot, A. D. (1951). War and the intelligence of youth. *Journal of Abnormal and Social Psychology, 46*, 596-597.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations. What IQ tests really measure. *Psychological Bulletin, 101*, 171-191.

Flynn, J. R. (1999). Searching for justice: the discovery of IQ gains over time. *American Psychologist, 54*, 5-20.

Gustafsson, J. E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg, *Advances in the psychology of human intelligence* Vol. 4, (pp. 35-71). Hillsdale, NJ: Erlbaum.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179-203.

Gustafsson, J.-E. (in press-a). On the hierarchical structure of ability and personality. To be included in a volume based on papers presented at the Second Spearman Seminar, Plymouth, 1997

Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28*(4), 407-434.

Harnqvist, K. (1968a). Relative changes in intelligence from 13-18. I. Background and methodology. *Scandinavian Journal of Psychology*, *9*, 50-64.

Harnqvist, K. (1968b). Relative changes in intelligence from 13-18. II. Results. *Scandinavian Journal of Psychology*, *9*, 50-64.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve. Intelligence and class structure in American life*. New York: The Free Press.

Husén, T. (1951). The influence of schooling upon IQ. *Theoria*, *17*, 61-88.

Husén, T., & Tuijnman, A. (1991). The contribution of formal schooling to the increase in intellectual capital. *Educational Researcher*, *20*(1), 17-25.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen D., Gintis, H., Heyns, B., & Mitchelson, S. (1972). *Inequality: A reassessment of the effects of family and schooling in America.* New York: Basic Books.

Lorge, L. L. (1945). Schooling makes a difference. *Teacher's College Record*, *46*, 483-492.

Lund, T., & Thrane, V. C. (1983). Schooling and intelligence: A methodological and longitudinal study. *Scandinavian Journal of Psychology*, *24*, 161-173.

Lubinski, D., & Humphreys, L. G. (1996). Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means. *Psychology, Public Policy, and Law*, *2*(2), 363-376.

Mårdberg, B., & Carlstedt, B. (1998). Swedish Enlistment Battery (SEB). Construct validity and latent variable estimation and profile prediction of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment*, *6*(2), 107-114.

Neisser, U., et al. (1996). Intelligence: knowns and unknowns. *American Psychologist*, *51*, 77-101.

OECD (2001) *Education at a glance. OECD indicators 2001 edition.* Paris: OECD.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

Winship, C., & Korenman, S. (1997). Does staying in school make you smarter? The effect of education on IQ in *The Bell Curve*. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.). *Intelligence, genes, & success. Scientists respond to the Bell Curve*. New York: Springer-Verlag, Inc.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

The evaluation of the Negotiated Curriculum Plan in South Australia

Liz Horrocks
Department of Education, Training and Employment *horrocks.elizabeth@saugov.sa.gov.au*

In the Department of Education, Training and Employment (DETE) in South Australia, IEP's for students with disabilities are called Negotiated Curriculum Plans (NCP's). There was a need to revise the NCP because of imminent policy changes in DETE. These include the rewriting of the Students with Disabilities Policy (1991) and the 2002 implementation of South Australian Curriculum Standards and Accountability Framework (SACSA) for all students.

Key stakeholders in the NCP process included parents, students, educators and service providers who were invited to take part in focus groups that were conducted in metropolitan and rural centres in South Australia. The purpose of the focus groups was to discover participants' understanding of the NCP, identification of its strengths and weaknesses as well as the essential elements of an NCP. Participants were also asked for suggestions for the format of the new NCP. The information from the focus groups was developed into a series of recommendations.

Negotiated Curriculum Plan, special education, parents, focus groups

INTRODUCTION

Philosophically, an Individualised Education Plan (IEP) for students with disabilities is the key document for providing curriculum access and participation and identifying special education services that underpin student achievement. IEP's have the potential to improve a student's education as they can specify the student's present level of performance, identify related services and state the goals that are to be achieved.

Purpose of the research

In the Department of Education, Training and Employment (DETE) in South Australia, IEP's for students with disabilities are called Negotiated Curriculum Plans (NCP's). There was a need to revise the NCP because of imminent policy changes in DETE. These include the rewriting of the Students with Disabilities Policy (1991) and the 2002 implementation of *South Australian Curriculum Standards and Accountability Framework* (SACSA) for all students. Currently, the NCP refers to *Statements and Profiles,* which is the present mandated curriculum. Therefore, when the new curriculum is implemented, the language of the NCP would be out of date. The current NCP had first been developed in South Australia in 1991, then revised in 1997. It was opportune at this time to review and evaluate the strengths and weaknesses of the NCP.

Background

The Negotiated Curriculum Plan (NCP), is a key implementation strategy of the Students with Disabilities Policy (1991), in DETE. Students eligible for support are those with a

'….physical, intellectual or sensory impairment and/or disabilities in communication and language. These students will have been referred for special education support by the guidance or speech pathology service' (Students with Disabilities Policy 1991). Students receiving support through the policy are required to have an NCP developed and documented at significant times in their schooling. This includes initial enrolment, at points of transition, when transferring from one school to another or when a guidance officer or speech pathologist refers the student for special education support.

Often schools review the NCP annually or more frequently on a need's basis. The purpose of the plan is to bring together the key stakeholders to negotiate the curriculum access and participation needs of the student. This can include parents/caregivers, the school principal or their delegate, the student, the teacher and appropriate special education personnel. DETE and non-DETE support services and agency personnel may also be involved. The resultant plan includes identified strengths, needs, agreed actions, review processes and related responsibilities. Issues of Aboriginality, low socioeconomic status, isolation, culture and gender are also taken into account.

Research questions

At the outset of this project, it needs to be made clear that the IEP that is being evaluated is the South Australian NCP, not IEPs in general. The questions asked of the focus group participants can be found on page 9 and relate to the participants' understanding of the purpose of the NCP, what are it's strengths and weaknesses and what should the essential elements of a new NCP be. The participants were also asked for suggestions for the format of the NCP. The information from the focus groups was developed into a series of recommendations from which the new NCP will be developed.

In considering what the future NCP may look like in South Australia, it is interesting to look at the history of educational planning for students with disabilities.

LITERATURE REVIEW

Historical perspectives

The concept of individualised planning in special education is not new. Over a hundred and thirty years ago Dr Edouard Seguin (1866) discussed individualising education for children with special needs. Dr Sequin wrote that: "*…the individuality of the children is to be secured, for respect of individuality is the first test of the fitness of a teacher…(individual planning) will secure the sanctity of true originality against the violent sameness of that most considerable part of education, the general training"* (p 26).

This was a plea to include individual planning in teacher training so that uniqueness of the learner could be accommodated and was acknowledgment of the fact that learners' needs differ. Approximately 100 years later, John Dewey's philosophy of American Education had a profound affect on educators. Dewey wrote that each child is '*entitled to equal opportunity of development of his own capacities, be they large or small in range…each has needs of his own, as significant to him as those of others are to them*"(cited in Abernathy 1959, p.254).

In recognition of this, educators began using a number of strategies to meet the individual needs of students. Schrag (1996) includes among these team teaching, practices for determining strengths and weaknesses and acknowledgment that learners have different learning styles. Dewey's ideas and approach to education have been given a fresh perspective by those who currently follow the paradigm of the constructivist theory (Goddard, 1997).

Individualised Education Plans

For the purposes of this paper, IEP's in the United States of America, Britain and South Australia were examined.

United States of America

In a briefing paper published by the National Information Centre for Children and Youth with Disabilities (1999), it was reported that Congress found that research into education in the last 20 years demonstrated the following. Effective education systems of the present and future should maintain high academic standards and clear performance goals for students with disabilities. These should be consistent with the standards and expectations for all students in the education system. As well, appropriate and effective strategies and methodologies should be put in place to enable maximum opportunities for students to achieve this. Therefore, the IDEA Act requires the IDEA team to determine access and support issues for students to successfully engage in the classroom curriculum.

The Vermont Education Department (USA) has created an IEP for students with disabilities and those who have difficulty meeting educational standards for that state. This plan uses "common sense" language, authentic assessments, holistic planning and uses the curriculum framework for goal setting and identification of standards (Giangreco, 1996). This format could be considered in relation to the data collected from the focus groups in this research, as the new curriculum in South Australia is standards-based. Webster and Came (1996) suggest that rather than the content driven IEP that appears to be unachievable, a process driven IEP would be preferable. Such an NCP would begin with description of need rather than a detailed catalogue of the student's difficulties.

Britain

In England and Wales currently, as in America, the IEP is a key component in accountability. Schools are expected to adopt them as a tool for setting out individual student's needs, ensuring teaching targets are appropriate and stated goals are delivered in the time set (Webster & Came, 1996). Besides accountability, IEP's in England are also a means of checking the way in which a school works for students with special educational needs (SEN). This includes how a school identifies, assesses, records, reports, communicates, co-ordinates, designs, plans, costs, allocates, delivers, reviews, monitors and evaluates. Teachers in England are not convinced of the value of an IEP in delivering the functions mentioned above because of the time it takes to put one together Webster and Came (1996). It would appear that too much is expected from one document. '*IEP's will prove useless as prompts for teaching and learning if they simply record masses of information and provide lists of teaching objectives. Decisions need to be made about who needs to know what, and for classroom practitioners and pupils themselves, how they set about managing learning for themselves'* (Webster & Came 1996, p8).

The Code of Practice

In the United Kingdom as a result of the Education Act (1993) schools are required to follow a national Code of Practice, which provides a range of structures and processes designed to support students with a wide range of learning needs. The Code has five stages of assessment and provision, which move from the individual classroom level, to the school level, to the involvement of external support agencies and finally to a formalised external assessment and statement process. The Code of Practice was revised in 2000 with the consultation phase due for completion in 2001 (Department for Education and Employment 2000). An IEP is required at Stages 2 and 3 of the Code of Practice and they are regarded as summary documents outlining practical plans that are based on past, present and future classroom activities, and of direct knowledge of the specific needs of the student (Cooper 1996).

Solity & Deavers (1999) argue that the proposed 60 million pounds that have been set aside for the development of IEP's in the United Kingdom in 2001, needs to be challenged. The reason for this is that such an investment carries with it the message that the current narrow practice should be maintained rather than an evaluation of its efficacy (Webster and Came 1996; Solity & Deavers 1999). In the IEP, they say, more importance is placed on where and how it happens rather than linking it to the teaching and learning process, its effectiveness and whether it changes children's learning outcomes, life expectations and experiences. Making South Australia's future NCP more effective by linking it to the teaching and learning processes within all key learning areas may help students to progress. If students with special learning needs are expected to progress, their learning will be interpreted differently and ultimately positive expectations will be conveyed to them. When all teaching proceeds on the basis of planned high expectations it will be reflected in higher student achievement (Solity & Deavers, 1999). Lucille Eber's Wraparound Process is useful to look at in relation to this.

The Wraparound Process

Wraparound is an approach to planning and implementing child and family centred services and supports. The process was designed for students with emotional problems (Eber & Nelson, 1997). However, there are elements that are appropriate to consider in terms of the future plan in South Australia. Thinking differently about current service structures, student needs, family partnerships and professional roles is the key in this process. Effective collaboration that reflects services rather than the disability that the child may have, is vital in this process. Eber's approach is to assess the problem, look at services or programs that are available and connect the child into those existing programs. Wraparound planning depends on blending natural supports with traditional services, for example, parent partners, community mentors and respite providers (Eber & Nelson, 1997). The key features of Wraparound are that:

- it must be community based
- services and supports must be tailored to meet the needs and identify the strengths of children and families, not to meet the needs of the system
- there is cultural competency building on the unique values, strengths, and social and racial characteristics of children
- there is family access, voice and ownership
- agencies must have access to flexible funding
- the process must be implemented on an interagency basis and owned by the community
- services must be unconditional
- outcomes must be measured.

This process is based on the principle of constructivism that builds on the unique values, strengths, and individual characteristics of children and honours the integrity of children as learners (Reid, Kurkjian & Carruthers 1994).

The members of the team who support the child are those in daily contact with the child and who have control over major resources or decisions. The extended team members are people who may come in to discuss resources, transitions or specific intervention. At an initial team meeting, a profile of life domains is developed that identifies: strengths, what has or is working for the child, perspectives and needs of the child. The profile also determines what needs immediate action and identifies what additional information is needed. One of the criticisms of the NCP is that often the meetings have too many members whose role is unclear. Eber's idea of members only attending a meeting if there is something to contribute, is sound, and may prevent too many people attending meetings unnecessarily. Determination of the need to be in attendance could be decided at the previous meeting.

Wraparound contains elements that include the strengths and needs of the student, lateral thinking, collaboration, family access, voice and ownership and only includes at meetings those

people directly connected to the current needs of the child. Comparison of the findings of the focus groups and these elements of the Wraparound Process will be made in a later section of this report.

Negotiated Curriculum Plans

In South Australia, students with disabilities have an NCP that documents curriculum access and participation needs and services that underpin and support learning. The reason for this is partly that Garth Boomer, Associate Director General of Education (Curriculum) in the early 90's advocated that curriculum for all students should be negotiated so that teachers and students could collaboratively develop an appropriate learning plan across the required areas of study. Such a philosophy enabled the focus for student learning to be within the required areas of study of the mandated curriculum and avoided skills based individual education plans.

The writers of the Students with Disabilities Policy (1991) adopted this philosophy. Two of the key people involved in the development of the original NCP offered the following anecdotal information to the researcher in the latter part of 2000 (Assistant Director, Disability and Professional Services DETE; Manager, Disability Services, North Group of Districts, DETE.).

> As part of the first draft of the policy, special educators talked about curriculum that was appropriate and negotiated. This term was loosely used until a decision was made not to use the word appropriate. In South Australia, there was a strong feeling amongst special educators that there should be a true negotiation between parents, educators and service providers, and that learning for students with disabilities should be part of the mandated curriculum. Hence, the name negotiated curriculum plan.

Students with disabilities at that time needed to be able to access curriculum that was describable, achievable, assessable and powerful (Bigge, 1989), as there was little documentation of student involvement in the key learning areas, little accountability and the curriculum did not necessarily empower students to excel. Therefore, the first NCP, designed after the introduction of the Students with Disabilities Policy (1991), included students with disabilities in the key learning areas through the documentation of access and participation needs using negotiated inclusive processes. The rationale for this policy is that: *"All young people in South Australia have the right to gain through the public schooling system, a broad balanced education that will prepare them for effective participation in society"* (Educating for the 21st Century, 1990). Among this group of young people are students with disabilities who may require support in order to access and participate in a broad balanced curriculum.

Relationship of problem to previous research

Much has been written about Individualised Education Plans IEP's, but there is little research on the NCP in South Australia, although some surveys have been conducted on its effectiveness. A search for articles on the NCP was made on ERIC, AEI, Kinetica and Family Databases, as well as various library databases (UNISA, Flinders, Barr Smith, SALSA). Checks were also made at the AEU site on the Internet.

The first of these surveys about the NCP was conducted in 1993 by the South Australian Institute of Teachers' (SAIT) Special Education Consultative Committee. Review of the implementation of the Students with Disabilities Policy by the union, which included a random survey of schools to determine the effectiveness of the NCP process. Data identified concerns about the number of people involved, the complexity of the document and the time taken to complete it (South Australian Institute of Teachers, 1993).

A second survey was conducted across all levels of schooling by personnel in the Targeted Populations Unit in the Education Department in 1993. Out of 50 principals surveyed, 39 responded. According to Chapman (1995), 60% of the participants stated that the NCP was useful because it assisted in:

- the facilitation of school structures
- the formation of parent/school partnerships
- transition planning
- identification of the need for ongoing resourcing
- familiarisation of the process.

Information on the impact of the NCP on classroom practice was not requested in this survey. This is crucial information regarding the effectiveness of the NCP for students with disabilities and is a question that was asked of the focus groups in the current study. Chapman's (1995) thesis related to South Australian policy and practice for students with disabilities. She indicated in her research, that although teachers had not changed their teaching practice as a result of the NCP process and did not feel that the process assisted them to develop learning goals, there was improved communication between parents and teachers and special measures for individual students were documented.

Horrocks (2000) also identified parent and principal concerns regarding the NCP in her work on grievance resolution in special education. Some of the issues that emerged were the need for:

- training in NCP processes for both parents and school personnel
- including parents in the NCP process and valuing their input
- regularly updating the NCP
- clarifying the purpose of the NCP.

However, parents and principals who were interviewed for this research agreed that although these were ways in which the NCP could be improved, the NCP was an important document for the education of students with disabilities.

Education plans for students with disabilities are important as they formalise access arrangements that enable participation in the curriculum and document services related to the student's needs. The literature suggests that IEP's can be narrow and prescriptive and not connected to classroom practice (Weist & Griel, 1995; Rodger, 1995; Webster & Came, 1996; Goddard, 1997; and Solity & Deavers, 1999). Therefore, notice must be taken of the literature so that the mistakes that others have experienced may be circumvented in the future NCP in South Australia. Eber's Wraparound Process blends natural supports with service delivery and elements of this process could combine well with a constructivist approach and emergent themes from the literature as well as findings from this research. The next section describes the procedures used to gather information from participants who were involved in the research.

RESEARCH PROCEDURES

Introduction

This research project incorporates both quantitative and qualitative aspects. Focus groups provide qualitative data that were used to quantify information that formed the recommendations for the revision of the existing negotiated curriculum plan. The four major stages of qualitative research outlined by Miles & Huberman (1994) were used in this research report and include the:

- design of the investigation
- collection of data
- analysis of data
- summarising and integrating the findings

Design of the investigation

Structured focus group interviewing was chosen as the primary method of data collection for this research. A range of key stakeholders was invited to take part. These included parents, educators, students and service providers involved in the NCP process. The questions were

reviewed and revised several times before deciding on the final set and were tested out on colleagues in the field first.

Data collection

The techniques for data collection in this research included:

- thorough planning
- testing the chosen data collection instrument
- ethics and professionalism
- thorough preparation (Punch, 1998).

Preliminary analysis of data occurred concurrently with data collection so that the two phases merged. The reasons for this was that there was the danger that the task of analysing large amounts of data became overwhelming and this may have resulted in the project not being completed or a reduced quality of work (Miles & Huberman, 1994). Data was collected during 33 focus group interviews.

Focus groups

Focus groups are considered to be a valuable tool for social researchers and other professionals regardless of their particular field of inquiry where collective testimonies and narratives have gained credence (Denzin & Lincoln 2000). The reason that qualitative techniques and in particular focus group interviews have taken time to be accepted is related to a reliance on numbers and the view that quantitative methods are more acceptable and reliable than qualitative methods (Krueger 1994). This reliance on data collected through empirical observations is termed positivism.

The term Individualised Education Plan was introduced in the United States of America with the passage of Public Law (PL) 94-142, the Education for all Handicapped Children Act of 1975. This act gave legal protection to teaching arrangements required by students with special needs (Webster & Came, 1996). In 1990, the Individuals with Disabilities Education Act (IDEA) (PL 101-476) not only retained the IEP but expanded its scope to include 16 year old students in transition from school. The term IEP is used to refer to a plan for curriculum, social adjustment, vocational education, physical education and adaptive behaviour related to special education and related services (Rodger 1995). In America, Congress stated that the original purposes of the IEP were to; provide an opportunity for communication among the parties regarding the IEP, encourage planning, assure an annual review and serve as an accountability tool (National Information Centre for Children and Youth with Disabilities, 1999). This reflected the standards and related accountability systems used by many American states have the potential to exclude students with disabilities.

The influence of feminist and post modernist ethnographic studies has also created renewed interest in focus groups as a research method where postmodernists find it is possible to minimise the distance between researcher and participant. Feminists have found that positivistic research limits the ability to explore the real life experiences and situations. On the other hand, focus groups have the potential to maximise this (Le Compte, Millroy & Priessle 1992 and Denzin & Lincoln, 2000).

The focus group interview is a collectivist rather than an individualistic research method that focuses on the multivocality of participants' attitudes, experiences and beliefs (Denzin & Lincoln 2000). It describes people in natural situations where an accurate description relies on people disclosing information in a comfortable environment.

What are focus groups?

Focus groups were considered to be a very appropriate way of gathering data for the purposes of this research because of the benefits of empowerment, interaction and group dynamics.

Powell & Single's (1996) definition of a focus group is: *'a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research'* (p 499).

Templeton (1994) defines the essential character of a focus group as *'a small temporary community, formed for the purpose of the collaborative enterprise of discovery. The assembly is based on some interest shared by the group.*' (p4).

Uses of focus groups

Stewart & Shamdasani (1990), summarise the more common uses of focus groups. Two of the uses they describe are relevant for this research. The first is to obtain general background information about the topic of the NCP and second, to learn how participants talk about it. Focus groups were chosen as a way of gathering insights and shared understandings relating to the use of the current NCP and to obtain a range of perspectives about the process and content of the future document.

Sampling

Sampling is necessary because you can't study everyone, everywhere doing everything (Miles & Huberman 1994). In qualitative research purposive sampling is often used where sampling is done in a deliberate way with a purpose or focus in mind. For the purposes of this research, informant sampling was carried out where several steps needed to be taken to locate information-rich participants (Punch 1998). To encourage people with information to share, to attend the focus groups, a memo about the review of the NCP was written was and sent to District Superintendents, Managers Support Services, Managers Disability Services, Managers Early Childhood Services, Managers Hearing Impairment Services, Principals and Directors. This memo informed district personnel of the review and consultation details. An advertisement was distributed to assist principals to advertise the process in their school newsletter. A student consent form was also distributed with the offer of transport for students if they wished to attend.

Method of ensuring confidentiality of information

Although a range of data about participants was collected eg age, gender, occupation, and level of schooling, names were not. No names were used in the subsequent write-up of the research. At the beginning of each of the focus groups, participants were assured of confidentiality regarding the information they gave and the researcher not falsifying or presenting results out of context, as set out by the Australian Health Ethics Committee of the National Health and Medical Research Council, (1996).

Permission to conduct the study

Permission was sought and granted to undertake this research from the Ethics committee of the Department of Education, Training and Employment and the Flinders University of South Australia. Arrangements were made regarding the storage of data.

Description of the participants

Participants were the key stakeholders in the negotiated curriculum planning process. The researcher wanted to obtain as broad a coverage of feedback from as many groups as possible. Parents or caregivers of students with disabilities are key people in the NCP process and hence were consulted. Too often, the most important person in the NCP process is forgotten and that is the student. Therefore, an opportunity to include students' views was considered to be of high priority. Unfortunately, only a few students were able to attend the focus groups. On enquiry it was discovered that in many instances, the school principal did not pass on the information about the focus groups to teachers, others said that that it was too difficult to

organise the students to come to the focus group venues. In hindsight, focus groups for students could have been conducted in their schools and information about the student focus groups sent out to special education teachers as well as their principals. However, some teachers said that they would like to involve their students, but they could not get to the focus groups. Consequently, written feedback was provided when the focus group questions were written down as a questionnaire.

Agencies such as the Crippled Children's Association, Autism Association, Down Syndrome Society and advocacy groups. Teachers, School Services Officers and Early Childhood Workers, Principals and Directors Guidance Officers and Speech Pathologists were also part of the consultation.

DATA ANALYSIS

Introduction

Miles & Huberman (1994) describe data passing through three interrelated stages: data reduction, data display and conclusion drawing and verification. These may not be sequential but must be seen as separate and adequately planned. The data from focus groups are the transcripts or notes of the group's interaction. This may be the only data gathering technique used in a study or combined with other quantitative or qualitative techniques (Punch 1998). In this research, data was scrutinised in two ways for the main themes, using NUD*IST software as well as manually searching the text for themes. The taped data from the focus groups was transcribed in a process that took six weeks. This was longer than the four weeks initially expected as there were difficulties with acoustics in some focus group venues. Several tapes had to be replayed to ensure accuracy of transcription.

Data reduction

Data was reduced in two ways. NUD*IST was the data management program used to store the large amount of data from the focus groups. This program was chosen because it offers an easy to use interface that handles research projects. NUD*IST searches documents and provides access to quotation, codes and memos (Punch, 1998). A NUD*ST tree was created that grouped the information and shaped the project. Search requests were made to collect data about a topic and then the searches were refined with more selection. For example, when requesting a search for the word "parents" it was evident that this produced an enormous amount of data without much definition. However, if the search was increased to 'parent communication' then there was less data but the information was more focused. Themes and topics changed as the data was explored.

To strengthen the results, transcripts were printed out, read and themes identified. Then a matrix display was set up using EXCEL. The number of times that a theme or topic was mentioned was counted and entered. With this number, percentages were worked out based on the number of participants the total group.

Matrix display and examination

A matrix display of the findings was constructed and used to present the results. The display needed to serve the function of providing answers to the research question being asked (Miles & Huberman 1994). It was used to present the data in a framework and most importantly, to show the patterns and explanations that became evident. For the purposes of this research, six columns divided the data into metropolitan and rural participants in the categories of educators, parents and service providers (see Table 4.0 below).

Table 4. Matrix display

Educators N=152		**Parents N = 45**		**Service providers N = 75**	
Metro N=79	Rural N=73	Metro N=25	Rural N=20	Metro N=58	Rural N=17

As only three students attended the focus groups, there was insufficient data to include them. However, a follow up project relating to students with disabilities and the negotiated curriculum plan, will be conducted by personnel in the Department of Education, Training and Employment in 2002.

Reliability

To maximise reliability, procedures that lead to consistent results were used no matter who moderated the focus groups were used. These included a set of structured questions that were used by all moderators. As well, training was provided and a pack with the same information given to all moderators.

Validity

Validity can be problematic if the data does not relate to what the researcher is describing. The validity of the research design lies in the extent to which it is actually capable of providing the information that it claims to provide (Dyer 1996). For the purposes of this research, content validity can be assessed, by examining the content of the focus group questions, against the intended purpose or aim of the research.

Bias

Researchers must be wary of intentional or unintended bias that can lead to misleading results. The research worker who moderated the focus groups had to be very mindful that personal bias did not lead to manipulation of the group or the answers the participants gave. The technique that was applied to obviate bias in this research was to form a rapport with the group as quickly as possible so that participants relaxed and once they were at ease began to talk. This meant that the moderator could let the group take over, giving less opportunity for bias.

Questions

The data was organised using the questions asked in the focus groups that were as follows:

1. what is the purpose of the negotiated curriculum plan (NCP) Should the NCP be a general summary of curriculum or an outcomes-based working program?
2. do you have any difficulties in the current NCP with process and/or content?
3. what do you believe should be the essential elements of an NCP?
4. what suggestions do you have for the format of the NCP? E.g. hardcopy, CD, website, other?

SUMMARISING THE FINDINGS

The feedback from the NCP focus groups was relevant and powerful because all the members had personal experience of the topic. Either the experience was as the person with the disability, the parent or caregiver who had unique knowledge of the student and are involved in the NCP process, the educators who work with the students or services or agencies who provide expertise and possibly additional funding. To strengthen the findings of the research, the data was matched to the key features found in a range of IEP's, Lucille Eber's Wraparound process and the IEP used in Vermont, USA. This information is presented in Table 5.2 after the advantages and limitations and before the recommendations.

Advantages of focus groups

There were distinct advantages in conducting focus groups that included:

- instant validity because issues could be checked before participants left the meeting
- the issues were real to the participants who were involved because they are key stakeholders in the NCP process
- focus groups gave participants an opportunity to discuss issues in an non-threatening environment
- the researcher was able to ascertain whether the participants understood the questions by using follow-up questions to probe the issues

The focus group consultation required considerable planning and organisation. Despite that, the groups were successful overall and participants were complimentary of the process. They appreciated the opportunity to voice their opinion regarding a process that has been frustrating for some.

Limitations of the Study

Focus groups do have limitations and in this research included:

1. fewer parents than educators participated in the focus groups, therefore, there could be a criticism that the information was 'system dominated'
2. there was difficulty in getting a larger number of students and parents to the focus groups
3. the researcher did not have absolute control over the data produced as it was inevitable that conversation veered off the topic at times, therefore, the moderator needed to be skillful and steer the conversation back when and where appropriate
4. focus groups are not fully confidential or anonymous because the members meet each other and some know each other
5. large amounts of data to be transcribed were quite expensive because it was time consuming. A person was employed to transcribe the data and the process took six weeks rather than four as was originally predicted

The issues that were raised were organised into the table of issues and recommendations below. These will guide the process, content and format of the future NCP. The data from the groups are particularly valuable and should be considered if the changes to the NCP are to be meaningful.

Comparison of desired key features of IEP's

To strengthen the findings from this research, features from a number of studies were reviewed in order to extract certain features. Although the new NCP must be compatible with the new curriculum in South Australia, it was found that they had certain features in common (see Table 5.4). These studies included the Vermont Family Centred IEP, a selection of national and international IEP's, Lucille Eber's Wraparound Process and the focus group findings.

The selection of IEP's included one from each of the following places: USA (California), UK (Hounslow), New Zealand, Western Australia (Wirrabirra and Forest Crescent), Queensland, and the South Australian Guidance Branch (Noarlunga House). There is also a selection of IEP's from Adelaide North Group of Districts and the McDonald Park Junior Primary School individual Education Plan. The South Australian IEP's, were given to the researcher by people who attended focus groups. The McDonald Park IEP was particularly interesting as the components matched the most desirable aspects of the work of Eber and the Vermont IEP as well as the findings of the South Australian focus groups.

From the table it can be seen that there is a strong connection between the findings of the research, key features of IEP's, the Vermont IEP and Lucille Eber's Wraparound Process. Eber's Wraparound is not an IEP, but an approach to planning and learning. However, characteristics of Eber's process should be considered for inclusion in the new NCP document. These include building on strengths and values of the student, measurable outcomes, family access, voice and ownership and encouraging community involvement where appropriate. Inclusion in the new NCP of the key features outlined in the table above should ensure that the resultant document is based on the feedback from the focus groups, key elements from a range of IEP's, and current literature. It should also serve to ensure that IEP problems encountered made elsewhere do not occur in South Australia.

Table 5.1 Comparison of desired key features of IEP's

Findings of the South Australian focus groups	**Key features from a selection of IEP's**	**Vermont IEP (Schrag 1996)**	**Lucille Eber's Wraparound Process**
strengths, needs, achievements and interests	strengths, needs and achievements	strengths and needs	build on the unique values, strengths, and social and racial characteristics of children
a single summary page	single page format	a new "non-form"- 3 blank columns - 1 for each key question	
there must be accountability for the NCP	evaluation accountability indicators of success	strategies to achieve a direct link between present levels of performance, services, goals and objectives	outcomes must be measured
the NCP must become more educationally meaningful	team includes parents, teachers, facilitators and community members	a holistic approach to enhance the IEP's educational relevance	the process must be implemented on an interagency basis and owned by the community
a clearly articulated process in jargon free, simple and specific language	clear processes in simple uncomplicated language	common jargon-free language	
include and value parents and students in the process	core and extended team to include the parents and student.	a creative problem solving process to promote participation by all team members	family access, voice and ownership
a clear, consistent process and transition checklist	clear checklist of processes	a checklist to ensure that legal requirements are satisfied.	

Recommendations

For the purposes of this report, the recommendations were set out below in Tables 5.2, 5.3 and 5.4 under the headings of content/format, process and support.

Table 5.2 Recommendations

Content/format	Recommendations
Accountability	The NCP needs to: be reviewed regularly; be outcomes-based; encourage teachers to take ownership for student learning
Content	The NCP needs to: be simplified with less paperwork; ensure compatibility with SACSA' determine long and short term goals; use language that is jargon free, practical, specific and simple; include a list of special centres, support services and other supports available; document teaching and management strategies that have worked or are not working for the student; reflect students' strengths, needs, achievements and interests; document what assessments will be used; document other relevant information; explore the interface with competency assessments to enhance the employability of students with disabilities
Format	The NCP needs to: be a simple learning needs framework allowing for the different needs of students; include physical access and curriculum assessments will be used be available as a CDROM
Expectations	In the NCP process: subject/classroom teachers must not leave the responsibility of the NCP to the special education teacher; realistic possibilities of what can and can't be done should be frankly discussed with parents
Reporting	Connections made between reporting, the curriculum and the NCP
Parent issues	The NCP should: use practical language with less jargon; include and value parents in the process; provide more information about how parents can help at home ; include the provision of a parent brochure or booklet
People from diverse linguistic and cultural backgrounds including Aboriginal and Anangu	There should be: translation of an explanation of the plan into other languages a pamphlet in different languages outlining the NCP process and available services more space in the NCP to talk about the student's background, abilities in their own countries and the psychological aspect of the student.

Table 5.3 Recommendations

Process	Recommendations
The NCP process	In the NCP process there should be: clarification of the process and its legal position consistent process information for educators, parents and students a simple, generic process checklist flexibility to meet the needs of different students provision by schools for release time to complete the process a complaint mechanism as a way of improving practice
Parent and student involvement	The NCP process should: include and inform parents and students ensure students are involved in the process encourage regular meetings with parents and establish good communication links
Secondary schools	Awareness of the NCP process should enable involvement in primary to secondary transition
Designation of management	The NCP should: clarify roles and responsibilities assist school personnel to prioritise resources
Accountability	The NCP should: address accountability assist in identifying curriculum priorities
Transition	The NCP should provide a: transition checklist information shared from site to site document roles and responsibilities in the transition process
People from diverse linguistic and cultural backgrounds including Aboriginal and Anangu	The NCP should recognise the need for: schools to be aware of the availability of Guidance Officers for people from diverse linguistic and cultural backgrounds a bilingual support worker to attend the NCP meetings for the purposes of translating and reporting parents to be asked if the student has any health issues or disabilities

Table 5.4 Recommendations

Support	Recommendations
Training and development	There is the opportunity for training of: principals. pre-service and in-service teachers, parents and agencies, teachers on the tribal lands in the North of South Australia and include information on the NCP in the induction of new teachers, on-line training
Resourcing	Schools need to allocate sufficient resources for the development of the NCP
People from diverse linguistic and cultural backgrounds including Aboriginal and Anangu	The NCP should recognise the need for: interpreters who are freely available and have a background in disability involvement of translators and MMALSA (advocates for people with a disability from other cultures) information about nationality group networks recognition in the NCP of the problems associated with torture and trauma.

Further research

Further research topics include:

1. more detailed research of the perspective of Aboriginal and Anangu parents, teachers, workers and students on the lands, in country towns and in the metropolitan area in relation to the NCP
2. the NCP and cultural perspectives
3. the effect on student learning of the new NCP
4. the South Australian Curriculum, Standards and Accountability Framework (SACSA), constructivism and the NCP
5. students and their involvement in the NCP
6. the professional development needs of parents and school councils in the NCP.

These topics will be prioritised and addressed by personnel in Student and Professional Services in 2001.

CONCLUSION

The use of focus groups as the method of data collection revealed insights and understandings relating to the current and future NCP as participants spoke candidly about the real experiences they had with it.

Three things are abundantly clear from the findings of this research. One is that the information from the focus groups will be pivotal in the formation of the future NCP in South Australia. Second is the value of the process of consulting in this way. Stakeholders appreciated being asked their opinion about current and future practice. The large degree of agreement about many issues was quite striking. The third is the number of issues that have been uncovered and will be explored.

Education plans for students with learning needs are key documents for providing curriculum access and participation, including special education services that underpin student achievement. It was important to listen to the voices from the field as well as others directly concerned in the NCP process. Focus groups were an economical and appropriate way of doing this. Not only was it important to listen, but also to consider the literature in relation to

education plans, so that the problems of narrowness and prescriptiveness inherent in the education plans of USA and UK are not repeated in South Australia.

This review has the potential to change the way that key stakeholders view and use the plan for students special learning needs, and shows that there is a considerable amount of agreement about the new direction. The new document should address student strengths and needs, accountability and parent and student involvement in a simple format that has clear, consistent processes. Cultural and linguistic issues and needs should also be taken into account. If these things are achieved, then the objective of this research has been reached. The NCP will have been reviewed and renewed in line with focus group feedback and current literature, incorporating aspects of student/family centred processes such as the Wraparound. The next stage of the NCP process is to design a functional, meaningful document for students with special learning needs in South Australia.

REFERENCES

Abernathy, G. L. (1959) *The idea of equality.* Richmond. VA: John Knox Press.

Bigge, J. (1988) Curriculum based instruction for special education students. Mountain View CA Mayfield.

Chapman, P. (1995) *Education of students with disabilities in South Australia: policy purpose and practice.* Thesis. University of South Australia, Underdale.

Daly, (1996). *Ethical Intersections: Health research, methods and researcher responsibilities*. Sydney: Allen and Unwin.

Denzin, N.K. & Lincoln, Y.S. (eds) (1994) *Handbook of qualitative research*. Thousand Oaks, CA:Sage.

Department for Education and Employment. (2000) *SEN Code of Practice on the Identification and Assessment of Pupils with Special Educational Needs.* Reference number: 0120/2000.

Department of Education, Training and Employment. (2000) *The South Australian Curriculum Standards and Accountability Framework* (Draft).

Eber, L. & Nelson, M. (1997) School-based wraparound planning: integrating services for students with emotional and behavioural needs. *American Orthopsychiatric Association* 67 (3), July.

Edelen-Smith, P. (1995) *Qualitative determinants for collaborative goal setting on the IEP*. Department of Special Education, University of Hawaii at Manoa.

Education Department of South Australia. (1990) *'Educating for the 21st Century. A Charter for Schooling.* Department of Education, Training and Employment. Adelaide. South Australia.

Education Department of South Australia. (July 1991) *Students with Disabilities Policy.* Department of Education, Training and Employment. Adelaide. South Australia.

Giangreco, M. (1996) *Vermont interdependent services team approach A guide to coordinating educational support services.* Paul H Brookes Publishing Co. Baltimore Maryland.

Giangreco, M., Cloninger, C., & Iverson, V. (1998) *Choosing outcomes and accommodations for children. A guide to educational planning for students with disabilities.* 2nd Ed. Paul H Brookes Publishing Co. Baltimore Maryland.

Goddard, A. (1997) The role of individual education plans/programmes in special education: A critique. *Support for learning.* Vol 12 No. 4.

Holman, H. (1993) Qualitative inquiry in medical research. *Journal of Clinical Epidemiology*. 46 (1) pp 29-36.

Horrocks, E. (2000) *Grievance resolution in special education: a descriptive evaluation of the issues.* Unpublished. Flinders University.

Krueger. R.A. (1988) *Focus groups: a practical guide for applied research*. London: Sage.

LeCompte, M.D.Millroy, W.L. & Priessle, J. (eds). (1992) *The Handbook of Qualitative Research.* San Diego, CA: Academic Press.

Margolis, H. & Truesdell, L. (1987) Do special education teachers use IEP's to guide instruction? *The Urban Review*. Vol. 19, No 3.

Miles, M.B. & Huberman, A.M. (1994) *Qualitative Data Analysis.* 2nd edn. Thousand Oaks: Sage.

National Information Centre for Children and Youth with Disabilities. (1999) *Briefing Paper. Individualised Education Programs*. LG2, 4th Edition, September.

Pearson, S. (2000) The relationship between school culture and IEP's. *British Journal of Special Education* Vol 27, No. 3 (September).

Poplin, M. S. (1995) Looking through other lenses and listening to other voices; stretching the boundaries of learning disabilities. *Journal of Learning Disabilities*. 28 (7), 392-8.

Powell R.A. & SingleH.M. (1996) Focus groups. *International Journal of Quality and Health Care* 8 (5): 499-504.

Punch, K.F. (1998) *Introduction to Social Research: Quantitative and qualitative approaches.* Sage Publications: Thousand Oaks.

Reid, K., Kurkjian. C., & Carruthers, S. (1994) Special Education Teachers interpret constuctivist teaching. *Remedial and Special Education*. Volume 15, Number 5, September, pages 267-280.

Rodger, S. (1995) Individual education plans revisited: A review of the literature. *International Journal of Disability, Development and Education*. Vol.42, No. 3, pp.221-239.

Ryan, A. (1994) *Targeted Populations. Unpublished survey*. Curriculum Division, Department for Education and Children's Services.

Schrag, J. (1996) *The IEP: Benefits, challenges and future directions.* Project FORUM. National Association of State Directors. US Department of Education.

Seguin, E. (1907) *Idiocy: and its treatment by the physiological method.* Albany, NY: Brandow Printing. (original work published 1866).

South Australian Institute of Teachers (1993) *Review of the implementation of the Students with Disabilities Policy.* SAIT, Adelaide, South Australia.

Stewart, D.W. & Shamdasani, P. (1990) *Focus Groups: Theory and Practice*. Newbury Park, CA: Sage.

Templeton, J., F. (1994) *The Focus Group*. (Revised Ed). McGraw-Hill. New York.

The Australian Health Ethics Committee of the National Health and Medical Research Council. (1995) *Ethical aspects of qualitative methods in health research*. Commonwealth of Australia.

Van Krayenoord, C., Elkins, J., Palmer, C., & Rickards, Field. (2000*) Literacy, Numeracy and Students with Disabilities. Centre for Literacy, Language Education Research.* Griffith University. Queensland.

Webster, A. & Came, F. (1996) Individual Education Plans, teaching and learning. *Special Children.* Questions Publishing: UK.

Weist, D.J. & Kriel, D.A. (1995) Transformational obstacles in special education *Journal of Learning Disabilities*. 28 (7). 399-407.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Some aspects of moral values among university students in three societies: Poland, Australia and the Philippines

J.J. Smolicz, D.M. Hudson and **M.J. Secombe**[1]
Graduate School of Education, Adelaide University

The paper reports the findings from a comparative study of moral values undertaken among university students from five universities in the Philippines, Poland and Australia. The aim of the study was to compare the way students brought up in different cultural traditions, but subjected to the homogenizing trends of globalization, evaluated what they deemed as the most appropriate solutions to some of the basic life dilemmas facing individuals today. The respondents were senior undergraduate and postgraduate students drawn from education and social science faculties in two independent Catholic and three secular state universities.

A total of 448 students participated in the study, with 80 to 100 students originating from each university. This paper presents a detailed analysis of one example from each of three categories of moral dilemmas, in which students were asked to state their position in relation to twelve issues including abortion, extra-marital affairs and stealing. In addition to providing concrete data on themselves, students indicated whether they believed the actions listed were wrong, acceptable or dependent on circumstances.

The students' responses to the issues showed a generally consistent position for each university along a value orientation continuum ranging from Traditional Christian to 'Post-modern Individualist'. The degree of support for Traditional Christian values varied, however, from issue to issue according to the university and country under investigation.

moral values; abortion; infidelity; stealing; tertiary students; international.

INTRODUCTION

As a result of global secular influences, it could be expected that responses to moral dilemmas would have become more uniform across different cultural and ethnic and religious groups. In order to ascertain any possible modification to or rejection of traditional moral values cross-culturally, the study sought the views of young people who were university students in three different countries and in institutions with different ideological profiles. The researchers had access to the collection of comparable data in Poland, the Philippines and Australia - countries very different in their regions, cultures and histories. It should be stressed that the study did not aim to take the form of a statistical analysis of variables to test a pre-established hypothesis which could lead to any predictions for a wider population.

[1] This comparative study was not funded by any organisation, but was the result of co-operative efforts among colleagues in the three countries concerned. The authors, in particular, would like to express their gratitude to Ms Monika Koniecko, Professor Elzbieta Halas, Dr Illuminado Nical, Ms Susana Manzon and Mr Robert de la Serna for the contribution they have made to the data collection and analysis of this paper.

The theoretical framework of the study is based on the humanistic sociological conceptualization of Znaniecki, in which the term 'moral values' refers to the meanings which group members are expected to accept as a guide to right and proper conduct and the way in which they should act in their daily life. Znaniecki (1963: 267-70) draws an important theoretical distinction between the ideological attitudes which individuals profess in terms of the religious and moral values of their group, and the tendencies to action which these same individuals reveal in the specific contexts of daily life.

UNIVERSITY PROFILES

Data were gathered from a total of 448 respondents, drawn from five different universities. The number in each institution ranged from 80 to 109 (80 in Leyte and Adelaide; 89 in Warsaw; 90 in UA&P and 109 in KUL).

In Poland, the student respondents were drawn from two universities. One of these can be regarded as firmly set within the Catholic tradition. In 1918 the Catholic Church established the Catholic University of Lublin (KUL) as a privately funded institution in a provincial centre in the east-central part of Poland. Over the period of Communist rule, it functioned as the bastion of Catholic religious and moral values in Poland. Today it is a medium-sized university with faculties reflecting its religious base: Theology, Canon and Civil Law, Christian Philosophy, Humanities, and Social Sciences as well as specialised institutes in Polish Church History, Higher Religious Education and Migration. Many of its students have gone to the priesthood and a variety of leadership positions in the Church and society at large.

The second Polish university which provided respondents for this study was Warsaw. A state institution, originally founded in the early nineteenth century, Warsaw is a very large university, with a full range of humanities, social science and natural science faculties. Its position in the heart of a capital city, renowned for its high culture in music, art, literature and the theatre, yet traditionally strongly linked to the Polish Catholic Church, has had an important influence on its ethos in providing access to moral values based in the Catholic tradition, as well as its more atheistic, anti-clerical counter-tradition.

Two other groups of respondents were drawn from universities in the Philippines. One set was provided by the University of Asia and the Pacific (UA&P), which is a private university in Manila, established by the Opus Dei order of the Catholic Church. Its ethos reflects the Opus Dei dedication to Catholic spirituality and emphasis on intellectual, political. educational and business leadership. It is best known for its high profile research specialisations in economics, political economy and business management, as well as its undergraduate College of Arts and Sciences. Its fee-paying structure and high entrance standards make it an élite institution, both socially and academically, and its total enrolment of around two thousand students is still relatively small.

The second Philippines institution from which respondents were drawn was the provincial state university at Tacloban on the island of Leyte in the eastern Visayas. The Leyte Institute of Technology (LIT) sees its role as providing for the practical educational needs of the surrounding community which has maintained a strong commitment to Catholic values reflecting the local community. It offers degrees in vocational disciplines, applied science and technology and teacher education, with a particular emphasis on post-graduate studies in education. At present time there are around seven thousand students enrolled at the Leyte Institute.

The two universities in the Philippines were chosen to provide contrasting communities within the overall Catholic ethos of the country. UA&P could be regarded as an intellectual centre closest to the teachings of the *magisterium* of the Church. Opus Dei is rightly regarded as a bastion of traditional Catholic teaching because it has scrupulously upheld the full and undiluted doctrine of the church. LIT, in contrast, is a comparatively small regional university,

chosen because it had no metropolitan glitter about it, was state-run and attracted a technically oriented clientele from financially and socially modest homes, in one of the poorest provinces of the country. The strong Catholic orientation of the Philippines was very apparent, however, in that seminars were normally begun with a prayer and Catholic feast days were regularly observed, even though it was a state university.

For the Australian section of the study, respondents were drawn from Adelaide University, which was established as a public institution in the late nineteenth century, the third oldest university in Australia. Situated in the capital city of the state of South Australia, it offers a wide range of degrees across the disciplines of science, agriculture, engineering, the professions, economics, music, humanities and social sciences. The student body currently numbers close to thirteen and a half thousand. Like most government-funded educational institutions in Australia, Adelaide University is strongly secular in ethos and has traditionally regarded religion and theology as sectarian pursuits, inappropriate for university students, while its philosophy department has been openly atheistic (Duncan and Leonard, 1974). The students participating in the study came from the Graduate School of Education.

Poland and the Philippines, the two predominantly Catholic countries investigated in this study, both inherited a strong Catholic tradition anchored in the post-Tridentum spirit. The respondents from the four Polish and Filipino universities who claimed to be Catholic (over 80%) would formally be expected to uphold the religious and moral values prescribed by their church and know that any deviation would be regarded as a sin. Evidence on which values were being rejected or regarded more relativistically and how frequently constitutes valuable information on changes taking place in moral values in countries which have long been regarded as strongholds of Catholicism.

METHOD

Student participants from the five universities were asked to complete a questionnaire on their religious and moral beliefs. For students in the Philippines and Australia the questionnaire was in English, but a Polish version was used for the respondents from Poland. Some questions were designed to gather concrete data concerning the respondents' background: gender, age, language use, level of parental education and religious affiliation. There was also a series of questions to gather cultural data i.e. these were concerned with the students' beliefs and moral perceptions and other attitudes in relation to the moral/religious convictions among them.

Responses to Specific Moral Issues

When presented with moral issues which epitomized contemporary moral dilemmas, participants were asked to indicate whether they agreed that action concerned was wrong: whether its moral status depended on the situation; or whether they considered the action not wrong. The range of responses available was:

- Considered Wrong
- Situation dependent
- Not wrong
- No opinion

The questions and responses were thus deliberately framed in terms of the maintenance of traditional moral values.

Data were initially summarized in frequency distributions for each university. Of the twelve situations presented to the respondents, one example from each category of issues will be discussed. viz. Abortion - classified as a 'Life and Death issue': Extra-Marital Affairs was

one issue from the category of 'Sex and Marriage': Stealing was one of two issues classified as 'Human (Social) Relations'.

Abortion: A Life and Death Issue

This issue proved important in the analysis because of the polarisation of data. There was a clear distinction between the percentage of responses which considered abortion to be wrong from the three Catholic oriented universities, including LIT, and the two more secular ones. Traditional moral values on abortion were strongly maintained on this issue especially at the three Catholic oriented universities: UA&P (93%), KUL(80%) and LIT (74%) while in Warsaw (49%) only half gave evidence of supporting these values. Even lower, the Adelaide figure of 25% who considered abortion to be wrong, could be seen to reflect a society where abortion is legally available, virtually 'on demand'.

In the case of responses rejecting traditional values, Adelaide respondents (44%) stood out as having the highest percentage of rejectors of traditional values. Warsaw (12%) is aligned with LIT on 15% in this response - an indication of much lower rejection among students who lived in predominantly Catholic societies, than the secular society of Adelaide which largely accepted abortion.

Abortion was an issue of debate, depending on the situation, for students from Adelaide (30%) and Warsaw (34%) rather than for the three Catholic-oriented universities LIT (4%), UA&P (6%) and KUL (14%). These students were upholding the traditional Catholic teachings. The situational response indicated that some students considered an issue was neither right nor wrong in itself. If they had doubts about it, they could choose this option to record individual responses .

Extra-Marital Affairs: A Sex and Marriage Issue

There was a marked convergence of responses in relation to this issue. The majority of students from all five universities agreed in maintaining traditional opposition to such actions: UA&P were unanimous on 99%. Warsaw(70%) and Adelaide (69%) showed a similar stance in opposing this issue. In this regard, Warsaw and Adelaide respondents almost matched the frequency levels from LIT at 73% and KUL at 85%. The 'ranking' of the university responses is similar to the previous issue, abortion, but there is general consensus.

The responses on the upholding of traditional moral values had their mirror image in the data showing rejection of such values. There was strong evidence of less acceptance for 'extra-marital affairs' at Warsaw (1%) and Adelaide (5%) where responses to this option were at a very low level, comparable to UA&P(1%) and KUL (1%). The respondents at LIT (14%) showed some approval of the issue but the number was relatively small.

The third response option, reflecting a more situational approach, was valuable in highlighting issues where there was a high degree of ambiguity or uncertainty. Low scores, in contrast, suggested that the respondents in the group concerned had made up their minds to be either upholders or rejectors of traditional morality. The relativistic view was comparatively high and shared by as many as 26% from Warsaw and 23% from Adelaide indicating the doubts some students expressed depending on the situation.

STEALING: A Human Relations Issue

The issue of 'stealing' was also striking in the way a large degree of consensus was displayed among students from all universities but the pattern of response differed markedly from those shown for the other two issues. 'Stealing' elicited the highest response of moral rectitude with over 70% of students in all universities and over 80% in three of them considering stealing as 'wrong'. 'Stealing' was also the issue which elicited the highest level of support for any traditional value among Warsaw students (87%). With a corresponding figure of 82% at KUL,

stealing seems to evoke a particularly strong sense of 'wrong' among Polish students. It is interesting to note that UA&P students, who were the highest upholders of moral values on virtually all other issues, were below Warsaw at 81% in the case of 'stealing'. Adelaide students with their dual Protestant and secularist orientations, also demonstrated greater acknowledgment of 'stealing' as 'wrong, than for any other issue (74%). At Adelaide 41% of respondents indicated that they belonged to one of the Protestant denominations whilst 30% claimed that they had no religion - a response which did not occur at all in any of the other groups of participants.

The remarkable unanimity of responses for this issue from students at all universities was also seen in the complementary rejection of traditional values, in that only very few claimed outright that 'stealing' was not wrong. Only LIT showed a relatively high proportion of those rejecting traditional values, amounting to 14%. In contrast, while on most moral issues Warsaw and Adelaide displayed a high degree of permissiveness, Adelaide showed no more than 4% of respondents, and Warsaw had no-one who condoned stealing. The figures for KUL and UA&P were minimal - 3 and 2%. Overall stealing was not acceptable as an issue.

The minimal number of students at Warsaw and KUL, UA&P and Adelaide who openly condoned stealing has to be considered alongside the proportion who expressed relativistic attitudes by indication that their judgement of whether stealing was 'wrong' or 'not wrong' would depend on the circumstances. Respondents from Adelaide (21%) and UA&P (17%) expressed some doubt in relation to stealing which was somewhat higher than Warsaw (12%) and KUL (16%). The figure for LIT was as low as 6% which reflects the pattern for all universities as seeing stealing as 'wrong'.

These findings suggest that moral issues in social relationships are not necessarily seen to be tied directly to Christian beliefs, despite their specific prohibition in the Ten Commandments. Findings raise questions such as : Is stealing related to established 'legal' rather than 'moral' codes? For some Protestants, what constitutes sin and restitution for sin becomes very much more a matter for the individual conscience and there is a greater tendency to adopt relativistic attitudes where what constitutes right and wrong depends on personal judgement of the situation.

CONCLUSIONS

The various responses to these three issues provide valuable information on the relative 'holding power' of Christian (mainly Catholic) beliefs in Poland and the Philippines when compared with mainly Protestant (or religiously indifferent) Australia. The responses also throw light on differences among the various moral values examined, in the extent of their support or rejection.

In the three predominantly Catholic universities (UA&P, LIT and KUL), traditional Christian values continued to be upheld for 'abortion' in the life and death issue. In the two universities where there has been a tradition of secularist values, there is a much greater rejection of traditional moral values, together with a substantial number of those with a relativistic orientation expressing an essentially ambiguous stance. One issue on which Warsaw and Adelaide join forces with the Catholic oriented University is on the issue of 'extra-marital affairs', a 'Sex and Marriage Issue', with about two third of responses upholding a traditional negative view of such practices. This particular value appears to hold no relationship to Christian religious belief or non-belief, assuming a virtually universalistic dimension.

The same situation apprears to hold for the 'social relationship' issue - stealing. Responses from students in all five universities favour the upholding of traditional moral values for this issue which is considered more frequently 'wrong' by Warsaw students than by respondents from all other universities. There is a minimal rejection of the traditional norms that condemn 'stealing'.

It is the rejection of certain moral propositions when the respondents openly admit that certain actions which are contrary to traditional morality are no longer 'wrong' which put them in conflict with the religious authority to which they officially subscribe. What is clear is that some of the traditional values are changing within the orbit of particular Christian denominations.

The results suggest that believers display a spirit of selectivity among moral values showing traditional disapproval for certain actions, while revealing a proclivity for greater permissiveness in relation to others.

The issue of abortion shows that in the Catholic oriented universities, there is little evidence of collapse of traditional values and that globalization trends have not seriously undermined other traditional values. The other two moral issues discussed, extra-marital affairs and stealing, where there is a greater convergence of responses, can be viewed as highlighting more universalistic values which appear to go beyond the confines of any particular religious belief.

REFERENCES

Andres, T. D. (1980) *Understanding Values.* Quezon City: New Day Publishers.

Blumer, H. (1939) *Critiques of Research in the Social Sciences I: Appraisal of Thomas and Znaniecki's 'The Polish Peasant in Europe and America'.* New York: Social Science Research Council.

Duncan, W.G.K. and Leonard, R.A. (1974) *The University of Adelaide.* Adelaide: Rigby.

Ossowska, M. (1985) *Normy Moralne: Proba Systematyzacji.* P.W.N. Warszawa.

Smolicz, J.J. (1997) In Search of a Multicultural Nation, in R. J. Watts and J.J.Smolicz (eds), *Cultural Democracy and Ethnic Pluralism: Multicultural and Multilingual Policies in Education.* Frankfurt am Main: Peter Lang.

Smolicz, J.J. (1999) *On Education and Culture.* (M.J. Secombe and J. Zajda, eds) Melbourne: James Nicholas Publishers.

Smolicz, J.J., Secombe, M.J. and Hudson, D.M., (2001) Family Collectivism and Minority Languages as Core Values of Culture among Ethnic Groups in Australia, *Journal of Multilingual and Multicultural Development,* vol. 22:2.

Znaniecki, F. (1963) *Cultural Sciences: Their Origin and Development.* Urbana: University of Illinois Press.

Znaniecki, F. (1969) *On Humanistic Sociology: Selected Papers* (Bierstedt, R. ed.). Chicago: University of Chicago Press.

Znaniecki, F. (1998) *Education and Social Change*. (Halas, E. ed) Frankfurt am Main: Peter Lang.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Small languages never die, they only fade away: The case of Welsh in Australia

Dr Arthur F. Hughes
Regency Institute *arthur.hughes@regency.tafe.sa.edu.au*

This paper is based on the proposition that the Welsh language is the core element of Welsh cultural identity. Drawing on quantitative and qualitative data, it presents a brief outline of those findings of a 1995 historico-sociological study of Welsh language maintenance and inter-generational transmission by a group of post-war Welsh migrants to Australia which bear directly on the school experiences of the second generation. The findings are presented against a brief background of related aspects of the current language situation in Wales.

Amongst the participants in this study, Welsh is confined almost exclusively to the home-domain and to the first generation. Loss of proficiency by the second generation was accelerated in almost every case by the primary and/or secondary school experience. Although attitudes towards maintenance were largely positive amongst both Welsh and non-Welsh speakers, they were seldom activated in terms of usage and support. Given the absence of structural support outside the home, the consequent failure of inter-generational transmission, and the reduction in the number of Welsh-speaking migrant arrivals, the prospects for the survival of the Welsh language in Australia are, at best, poor.

INTRODUCTION

Historical and sociological literature on the language and cultural maintenance patterns of the more culturally distinctive non-English speaking communities in Australia abounds. Although the composite ethnic groups which the term 'British' embraces have by no means been neglected, the main focus has been on the Scots (eg Prentis, 1983; Richards, 1988) and the Irish (O'Farrell, 1986). Of the British migrant population in Australia, only the Welsh have been almost entirely overlooked.[1] The silence is broken only by the occasional dismissive comment. Thus Campion at a 1980 colloquium on 'The Celts in Australia' remarked, not entirely tongue in cheek :

> And the history of Australia may be stated simply: the English made the laws, the Scots made the money, the Irish made the songs. The Welsh? They contributed little to Australian history. Billy Hughes was not a Welshman as Donald Horne's recent book showed.(Campion,1980:1)

One cannot of course entirely disagree. Even with due disregard to the ideology of 'contributionism' which informs many studies of ethnic groups, it must be said that the impact of the Welsh, has been minimal. This is largely a consequence of their very small numerical presence, an invisibility further accentuated by the use, almost throughout the nineteenth century of the ubiquitous 'England and Wales' in official records.[2] To this must be added that amongst the Welsh-born, Welsh speakers are further concealed both by their very small numbers and, until recently, a dearth of information on the language proficiency of migrant arrivals. This lacuna in studies of the migration and settlement of the constituent British ethnic

[1] An exception would be Lloyd, (1988). This is, however, largely a historical account of Welsh settlement.
[2] Particularly at the British end of the migrant passage.

groups is rather curious since the Welsh are arguably the most culturally and linguistically separate.

This paper attempts a summary of some of the key findings of a 1995 historico-sociological study of Welsh language and cultural maintenance and inter-generational transmission by a group of first and second generation post-war migrants from Wales to Australia in so far as they relate to the school experience. For reasons of space and the theoretical assumption which underlies the study that the Welsh language is the core component of Welsh cultural identity, the focus here is on language rather than on other aspects of Welsh cultural life in Australia. Given recent crucial developments affecting the prospects for the future survival of Welsh in its homeland, the findings are presented against a brief background of its current condition in Wales. Even more than with the larger, and therefore [perhaps] more secure languages,[3] the fate of small languages in migrant societies is inextricably tied to the linguistic *status quo* in the country of origin.

THEORETICAL ASSUMPTIONS AND METHODS OF RESEARCH

The theory of culture in which the study is grounded was first developed by Znaniecki (1969) and subsequently refined by Smolicz (1979) for the study of ethnic cultures. As one of the mainstays of humanistic sociological theory it need not be outlined in any detail here. Two of its fundamental axioms, however, directly apply :

- the assumption that cultural and social phenomena can only be fully understood if they are studied from the viewpoints of the participants. This means the researcher must try to view the social world as it exists for those under investigation through a process of what Znaniecki (1969) calls 'imaginative reconstruction'. To this end subjects of study are encouraged to talk and/or write freely about their thoughts, feelings, aspirations and experiences. In the present context this theoretical stance translates into the use of oral and written memoirs as tools of research into the cultural and linguistic experiences of Welsh migrants in Australia.

- the notion that within the range of distinctive elements which constitute the culture of a group there are definitive or core values (Smolicz, 1981), the maintenance of which is the *sine qua non* of the group's survival as a separate cultural entity. Amongst these quintessential markers of identity, language is usually (but not always) the most manifest – it is certainly the most readily identifiable and is, in the most literal sense, the group's public voice and face to the world. For Welsh-speakers, and increasingly for non-Welsh speakers, it is the Welsh language that has become the rallying point of cultural survival, the last ditch of 'authentic'[4] C*ymreictod* (Welshness).

- Multiple research methods were used. These, for the most part were *qualitative* [oral and written memoirs, participant observation, informal interviews and documentary evidence] but three separate self-administered, Welsh-English bilingual questionnaires were designed in order to obtain concrete[5] fact profiles of the respondents and identifiable data on their evaluation and activation of Welsh cultural values. It must be stressed however that the research methods used were complementary rather than competing means of data gathering.

---

[3] Though it by no means follows. Size is no guarantee of linguistic vitality. See P.Nelde et al., (1996)..

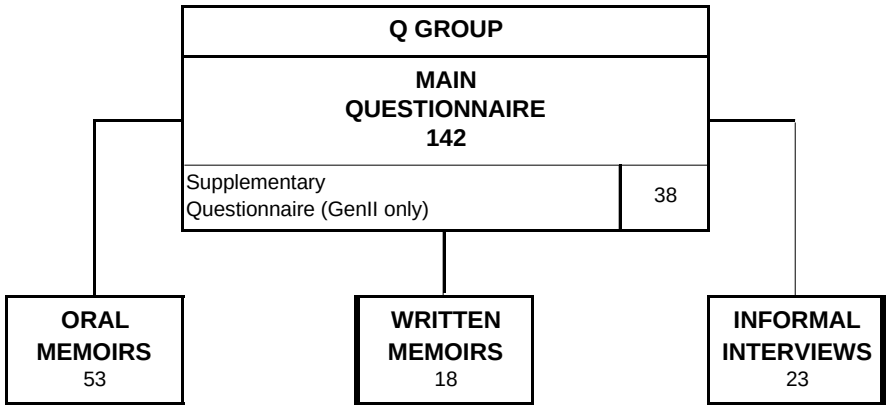
[4] Fishman (1966) makes the distinction between 'authentic' or 'primordial' ethnicity and its 'non-authentic' or 'transmuted' expression. In language-centred cultures the latter refers to the residual cultural values that are left behind after the demise of the language.

[5] 'Concrete' facts here refers to personal details on family background, education etc.

THE PARTICIPANTS

Sources of data and participant numbers are shown in Table 1.

Table 1. Source of Data: Q (Questionnaire) and NQ (Non-Questionnaire) Groups



TOTALS	
Main questionnaire (Q group only)	142
Supplementary questionnaires (Q group Gen II only)	38
Oral memoirs (Q group and NQ group)	65
Written memoirs (Q group and NQ group)	23
Informal interviews (Q group and NQ group)	36

The 167 participants in the study fell into two groups:

1. The primary or main group of respondents [Q group] comprising 104 first and 38 second generation[6] Welsh migrants to Australia all of whom completed a 28 page questionnaire. Within this group some 45% of the first generation and 19% of the second claimed to be Welsh-speakers. A supplementary questionnaire was also completed by the 38 second generation respondents.

2. A secondary group of 25 respondents [NQ Group] comprising a) non-permanent Welsh residents, b) key informants with specialist knowledge of the Welsh-Australian connection and c) returned Welsh migrants. All had direct experience or knowledge of Welsh cultural life in Australia.

The point should be made that the subjects were not chosen by the use of any mathematical random sampling procedure and the results were interpreted accordingly. No claim was therefore made that the 142 Welsh migrants who were the main subjects of investigation were, in any sense, representative of the totality of Welsh settlers in Australia. Conclusions reached, based as they were on the reconstructed cultural experiences of the participants were

[6] First generation {Gen1) is defined as adult Welsh migrants who arrived in Australia after the age of 12. Second generation comprises foreign-born children of Gen 1 who arrived in Australia before the age of 12 (Gen 2a) and Australian-born children of Gen 1 (Gen 2b).

exploratory rather than predictive, indicative rather than representative. This said, it must be pointed out that in a non-scientifically defined sense the respondents were 'randomly' recruited rather than selectively chosen. Most were chance contacts within diffuse kinship and friendship networks, including Welsh Societies and Nonconformist chapel[7] congregations throughout Australia.

THE CURRENT STATE OF THE WELSH LANGUAGE IN WALES

Needless to say the language situation in Wales is extremely complex and in some respects almost resistant to analysis. Only a brief survey of recent developments which relate directly or indirectly to the findings presented in this paper is possible here.

According to the 1991 census some 523, 319 or 18.7 % of the 2,798,500 population of Wales are Welsh-speaking. Undoubtedly in recent years Welsh has been institutionalised[8] and legitimated[9] within the boundaries of Wales more than any other minority language within the European Union. One significant result was that the 1991 census showed the first reduction in the decline of Welsh-speakers in almost a century. Moreover the most rapid growth was to be found in the 3-15 age cohort. This apparent turnaround in the fortunes of the language is obviously due to a variety of reasons including a militant civil disobedience campaign that *inter alia* saw the setting up in 1982 (Jones, 2000) of a Welsh-language TV channel. [10] There are also by now more than 50 Welsh-language local newspapers. High on the list of explanations for this remarkable development has been the establishment since the 1960s of a range of nursery, primary and secondary Welsh-medium schools throughout Wales but particularly in the non-Welsh speaking areas.[11] The latter point means that the majority of pupils in these Welsh-medium schools come from English-speaking home backgrounds.(Hughes, 2001)) There are also bilingual schools in the predominantly Welsh-speaking areas. The cumulative effect of what has been nothing less than a linguistic *Risorgimento* is that the ability to speak Welsh is now a requirement or at least an important asset for advancement within the public sector in Wales – a far cry from the conclusion of the Royal Commission Report into the State of Education in Wales, 1847:

> The Welsh language is a vast drawback to Wales and a manifold barrier to the moral progress and commercial prosperity of the people……..The evil of the Welsh language is obviously and fearfully great……it distorts the truth, favours fraud and abets perjury.[12]

SOME FINDINGS RELATED TO THE SCHOOL EXPERIENCES OF THE SECOND GENERATION

It should be pointed out that most of the participants arrived in Australia prior to the revival of the Welsh language briefly delineated above. They came for the most part in the 50s and 60s from a Wales in which the status of the language was still very low. It was essentially a private and domestic *patois* with little of the social and/or public standing it currently enjoys. Speaking Welsh in public settings was at that time almost a political act. This is an important link with the fate of the language in Australia in the post war years. It certainly bears crucially

---

[7] Welsh Nonconformist churches [Methodist, Presbyterian, Baptist] are called 'chapels'

[8] For instance the Welsh Language Act of 1993 repealed all previous legislation pertaining to the language, including the Acts of Union of 1536 and 1542 which proscribed its use The 1993 Act placed Welsh on a basis of 'equality' with English in the courts and in all domains of public administration. To ensure the effectiveness of this legislation a Welsh Language Board with statutory authority was set up This authority includes the right to demand that bodies within the public sector put forward practical schemes to maintain and promote the Welsh language. Needless to say the recent devolution of government in the UK and the establishment of the Welsh Assembly will only strengthen the force of these legislative measures.(Jones,2000; Welsh Language Board, 1995)

[9] Legitimation here is understood to mean formal recognition by the state, with a resulting increase in official use and raising of status. For a fuller discussion of the concept see May (2000).

[10] Wales now has 2 Welsh language TV channels including S4C Digital which is also broadcast to other areas of the UK .[with subtitles]

[11] By Welsh-medium is meant schools in which almost all subjects are taught through the medium of Welsh. Their growth has been remarkable. By the early 90s the nos were as follows; nursery [617]; primary [538]; and secondary [58].[May, 2000, 109:10]

[12] 'Report of the Commissioners of Enquiry into the State of Education in Wales, 1847' cited in Grillo (1989).

on the attitudes towards its maintenance and transmission amongst the subjects of the study under discussion.

Certainly, despite very limited opportunities to speak Welsh outside - and frequently inside[13] - the family home, the evidence of the study was that the adult first-generation speakers retained fluency almost indefinitely after arrival in Australia. The great majority of speakers [94%] and non-speakers [68%] also thought knowledge of the language crucial to the maintenance and transmission of a Welsh cultural identity. Only 8% of all respondents thought it of no importance that a Welsh person be able to speak Welsh. The following comment was typical:

> Why maintain the language? Rather a silly question if you ask me. Who wouldn't want to keep his own language? To me it's as important to hold on to it here in Australia as anywhere else.[14] [Oral memoir extract, Q Group; translated from the Welsh.]

Yet this typical positive attitude towards retention was seldom translated into active support. For instance, only 34% of speakers and 14% of non-speakers condoned the speaking of Welsh in public in the presence of or even within the hearing of English monolinguals. This attitude was prevalent even amongst Welsh community leaders. Speaking of his successor, the retired Welsh-speaking minister of one of the Welsh chapels[15] commented:

> We would often be coming out of the church together to mingle with the congregation in the foyer and to my great great annoyance he would insist on speaking Welsh to me. It didn't seem to bother him but it embarrassed me terribly I tell you. [Interview extract, Q Group; translated from the Welsh]

This disapproval of the public use of Welsh is of course related in part to issues of politeness but, more than with most other languages in Australia, has its roots in the historical unequal status of Welsh and English. It is also reflective of attitudes which have witnessed the very rapid decline of the language in Wales during the 19th and 20th century.

It bears also on many of the problems attending on transmission in Australia. Here again the disparity between attitude and actual language behaviour was very evident amongst the participants in the study. Almost all the parents in Q group [93%] thought it 'very important' to maintain and transmit the Welsh language to their children but very few had taken any steps towards its realisation. The consensus was that any effort was futile in the absence of institutional support outside the family home. The two surviving Welsh chapels[16] had few Welsh-language services,[17] proceedings in all the Welsh Societies were held in English, and there were no Welsh-language ethnic schools. Even within homes where both parents were Welsh-speaking the situation was equally bleak. Language loss by the second generation within a few years of arrival was almost universal. In almost every instance it was accelerated by entry to the first primary or secondary school in Australia (see Table 2).

At that stage also, that is when the children could no longer [or in some cases refused to] communicate with their parents in Welsh, the language of the home switched to English. A fundamental consideration here is that all the Welsh-speaking parents were also bilingual. Unlike many speakers of other ethnic languages in Australia they are free to choose between their own language and English when speaking to each other or to their children. *Speaking Welsh within and outside the family home becomes therefore an act of will not of necessity.* The latter point is crucial to maintenance and transmission. There is simply no *need* to speak

---

[13] As in the case of mixed linguistic marriages.

[14] First-generation Welsh-speaker

[15] In which services were bilingual.

[16] One in Victoria and one in New South Wales. In the 1860s there were 29 in Victoria alone. (Y Beirniad, 1873)

[17] Once a month on average.

Welsh on any occasion or in any domain in Australia. This means the second generation is seldom exposed even to the sounds of the Welsh language.

Table 2: Welsh Proficiency Levels of Second Generation of 'Q' (Questionnaire) Group at Start and End of First School Entered in Australia (N = 38)

LANGUAGE SKILLS	PROFICIENCY LEVEL (WELSH)																			
	VERY WELL				WELL				LITTLE				NONE				TOTALS			
	Start		End		Start		End		Start		End		Start		End		Start		End	
	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%
Understanding	10	26	5	13	3	8	1	3	5	13	4	11	20	53	28	73	38	100	38	100
Speaking	10	26	2	5	2	5	0	0	5	13	7	19	21	56	29	76	38	100	38	100
Reading	6	16	2	5	3	8	0	0	6	16	5	13	23	60	31	82	38	100	38	100
Writing	6	16	2	5	4	10	0	0	4	10	5	13	24	64	31	82	38	100	38	100

Interestingly a frequent complaint of the second generation was the failure of the school in Australia to recognise their English language difficulties:[18]

> When I first arrived I could only speak a few words of English. It took me about two years to settle down. I was what you might call a good student back in Wales but in Australia I was considered something of a slow learner because my English was so poor. The teachers somehow didn't understand that Welsh was not a dialect but a Different language and that it was my first and only language. As I came from Britain it was just assumed that English was my mother tongue. [Extract from Oral Memoir; Q group, second generation]

One Welsh-speaking family, aware after a number of years in Australia of the futility of any attempt to transmit the language to their children took the drastic step of returning to Wales. Table 3 illustrates the language development of one of their two children:

CONCLUDING COMMENT

Undoubtedly the Welsh language was highly valued by the majority of the subjects of the study discussed. Speakers and non-speakers saw it as the most important component of Welsh cultural identity in Wales and Australia. For the great majority it was, however, only peripheral to their lives. It was the natural medium of communication in only a very few households and was seldom if ever the language of inter-generational exchange. It was also infrequently used or heard outside the family home. When the children of Welsh-speaking parents started school in Australia they became English monolinguals within one or two years. This was the experience of all the Welsh-speaking families irrespective of parental attitudes and/or to a lesser extent the age at which the children entered school. Parents, anxious not to impede their children's academic progress, accommodated by switching to English as the sole medium of communication with them and, in some instances, with each other.

In Australia Welsh will probably retain its symbolic value as the prime marker of Welsh cultural identity. As such it will not die, but having ceased to be the language spoken at the family dinner table it will not be transmitted to future generations. Given also the unlikely prospect of Welsh-speaking migrant arrivals again reaching post war numbers, it will most probably fade away.

[18] Some of whom came from areas of Wales where even in the 60s English was seldom heard at school or in the street.

Table 3: Longitudinal case-study of the language development of a second-generation respondent [Q group]

Age	Speaking Proficiency	Explanation	Biographical Data
4	Monoglot Welsh	Native-speaker of Welsh.No knowledge of English.	Arrives in Australia [1976]
5	Welsh-dominant bilingual	Speaks some English but Welsh is his natural language of use. Hears only Welsh at home, speaks only Welsh to his parents.	Exposed to English via TV and visits by English-speaking cousins and friends [1976-8]
6	Balanced bilingual	Near equal facility in Welsh and English. His speech marked by frequent code-switching	Starts primary school [1978]
8	English-dominant bilingual	Speaks Welsh with difficulty and reluctantly. Feels more comfortable with English. Parents continue to speak Welsh to each other but increasingly in English to him and his younger brother.	Has attended primary school for 2 years.[1980]
10	Monoglot English [speech]; Receptive bilingual [understanding ability only of English]	No longer speaks Welsh to parents or friends. Parents speak Welsh to each other but English only to their two children	[1982]
12	Monoglot English	Has very little understanding of Welsh. His brother is also monolingual English	Enters secondary school in Australia
		Family returns to Wales [1984]	
13	English-dominant 'bilingual' [understands Welsh]	Speaks some Welsh but with difficulty to grandparents and family friends.	Enters secondary school in Wales in which Welsh is the medium of instruction in most subjects
16	Balanced bilingual	Speaks Welsh and English and is fully literate in both. Younger brother now also speaks. Welsh is the language of the home	Completes GCE 'O' level [including Welsh and English] Wins school prize in Welsh literature.[1988]

BIBLIOGRAPHY

Campion, E. et al (eds), (1980), '*Celts in Australia:Imagination and Identity'* Colloquium Papers, March.

Fishman, J. A. and Nahirny, V.C. (1966)'Organisational Leadership Interest in Language Maintenance'. In J.A.Fishman (ed.), *Language Loyalty in the United States',* The Hague: Mouton & Co.

Grillo, R.D., (1989), *Dominant Languages,* Cambridge: Cambridge University Press.

Hughes, A. F, (2001), 'Welsh'. In J. Jupp, *The Australian People: An Encyclopedia of the Nation, Its People and Their Origin,* Cambridge : Cambridge University Press.

Jones, G. (2000), *History and Status of the Welsh Language,* Online ; accessed 13/11/01. http://users.comlab.ox.ac.uk/geraint.jones/about.welsh/

Lloyd, L., (1988), *Australians from Wales,* Caernarfon :Gwynedd Archives.

May, S., (2000), 'Accommodating and Resisting Minority Language Policy: The case of Wales', *International Journal of Bilingual Education and Bilingualism,* 3 [2] 101-128.

Nelde, P. et al, (1996), *Euromosaic:The Production and Reproduction of the Minority Language Groups In the European Union,* Luxembourg:Office for Official Publications of the European Communities.

O'Farrell, (1987), *The Irish in Australia,* Kensington NSW: University Press.

Prentis, M. D., (1987), *The Scottish in Australia,* Melbourne: AE Press.

Richards, E. (1988), 'Australia and the Scottish Diaspora". In J. Hardy (ed.), *Stories of Australian Migration,* Kensington : New South Wales University Press.

Smolicz, J.J. (1979), *Culture and Education in a Plural Society* ,Canberra : Curriculum Development Centre.

Smolicz, J.J.(1981), 'Core Values and Cultural Identity". *Ethnic and Racial Studies,* 4, 75-90

Znaniecki, F. (1969), *On Humanistic Sociology,* Chicago : University of Chicago Press.

Y Beirniad, 1873, (X1V), 338-347. [Welsh Language Journal].

Web mathematics anyone?

Esther Yook Kin Loong
University of South Australia *looey001@students.unisa.edu.au*

The World Wide Web has had an impact on many areas of teaching and learning. Mathematics teaching however has only recently begun to utilise and develop this educational resource. This paper outlines a research program, which aims to uncover the extent the Internet, in particular the World Wide Web, is being used for High School mathematics education. The program includes searching out discernible Web-based teaching strategies and examining their impact on mathematics teaching and learning attitudes and achievements. Of particular interest is the extent to which deployment of the Web in mathematics teaching might increase student interest in mathematics. The first step in this process is to develop a preliminary typology of mathematical elements on the Web. The nature of these elements, their categorisation and their possible roles in the teaching and learning of mathematics are discussed.

mathematics, World Wide Web, typology, mathematics web sites, exemplars

INTRODUCTION

In this era of globalisation and information technology, educators have been urged to equip themselves and their students with skills that will enable them to successfully move forward into the future (Williams and Price, 1999). Doubtless, the next two decades will see new directions in education as "information technology paradigms change and organizational focus shift from on-site schools to networks of educational service bonded by web-oriented coursework and integrated delivery systems" (Miller and Miller, 2001). The wealth of information, activities and opportunities available on the Internet has caused many to advocate its use as a resource tool for teaching and learning (Barron and Ivers, 1998, Skomars, 1998; Butler, 1998). The World Wide Web, in particular, is a flexible environment for lifelong learning and has the potential to change the way curriculum is implemented and the way teachers teach (Stager, 1996; Quinlan,1997). The uptake of the World Wide Web for educational use, however, is relatively slow in the field of Mathematics and there is a scarcity of literature in this area.

Although use of computers in mathematics education is not new (Hansen and Zweng, 1984), teacher practice in the use of the Internet for mathematics education is at best in its infancy. Despite this, there are a considerable number of web sites for mathematics on the World Wide Web. Among these are active sites maintained by reputable organisations, universities and schools. The sites offer information, activities, games and tasks that can be used to enhance Mathematics education from kindergarten to high school. But to what extent have teachers used them and in what ways are they being used?

For Mathematics to be meaningful, students need a varied and adequate range of examples and a global structure so that misconceptions and perception difficulties can be reduced. Pereira-Mendoza and Kaur (1999) suggested that teaching should focus on relational understanding and on providing a wider and more varied context for the application of mathematics. This call is not very different from the one made more than a decade ago by

Fuys, Geddes et al. (1988) for teachers to look beyond a page of text because of its limitations for manipulative exploration and discovery learning. It has been found that when confronted with mathematical tasks, students tend to think in associative ways rather than constructive ways and they tend to have compartmentalised knowledge structures and are reluctant to relate the concepts unless forced to do so by the problem statement (Evans, 1991).

Numerous types of software programs have been used for the teaching and learning of Mathematics since the 1980's (Hansen and Zweng, 1984). The late 1990's have seen newer and more sophisticated mathematical software, such as *Derive, Geometer's Sketchpad, Mathematica* and *Maple,* being used in schools as well as in institutions of higher learning. While these software packages are without doubt useful and effective in their own ways, their educational use is restricted by licensing procedures and costs as well as their availability in students' homes. In contrast most mathematical activities on the Web are freely available for educational use (Barron and Ivers, 1998; Skomars,1998), and are a likely alternative for use in the mathematics classroom. At the time of writing, the author has identified more than 50 active web sites for mathematics teaching and learning, containing numerous associated web pages of colourful, animated and/ or interactive tasks and activities. Recent developments have shown that there are now available Web versions of *Mathematica* and *Geometer's Sketchpad.* Web*Mathematica* is a new technology that enables *Mathematica* functionality over the Web. It is currently being trialed in Denmark by the Danish Ministry of Education for use in 24 high schools. The courseware integrated with web*Mathematica* allows students to complete interactive lessons online and to submit their homework electronically upon completion through an intranet link (Wolfram Research, 2001).

The World Wide Web, with its myriad of information in written, audio and visual form is particularly suited to classroom use as a resource-based tool for a subject like Language Studies (Sexton, 2001). Characteristics of the Web such as authenticity, topicality, quantity, variety, hypertext links and searchability make it relevant to foreign language learning. What features in the Web would make it particularly relevant to Mathematics education?

An examination of these and other questions relating to the use of the Web in mathematics education requires an extensive research program. Such a program would identify and categorise useful "sites", provide a mapping of relevant "sites" to topics in the curriculum, describe effective teaching strategies that deploy such "sites", study their educational effectiveness and describe impacts on student behavior and motivation. The following outlines the stages of a research study that examines some aspects of the above program.

THE STUDY

The focus of this study is to determine the extent teachers are using the Web for high school mathematics teaching and learning, the teaching strategies they adopt and the impacts on students.

Design

This study will be conducted in 3 linked stages. Action in Stage 2 will be contingent to the results in Stage 1.

To determine the extent the Web is being used in the teaching of high school Mathematics, a broad paper-based survey will be conducted on mathematics teachers teaching Years 10 -12 in selected schools around South Australia. A stratified random sample of public and private

schools in urban as well as country areas will be taken. An online survey will be open to any individual who is a secondary Mathematics teacher and is interested to participate.

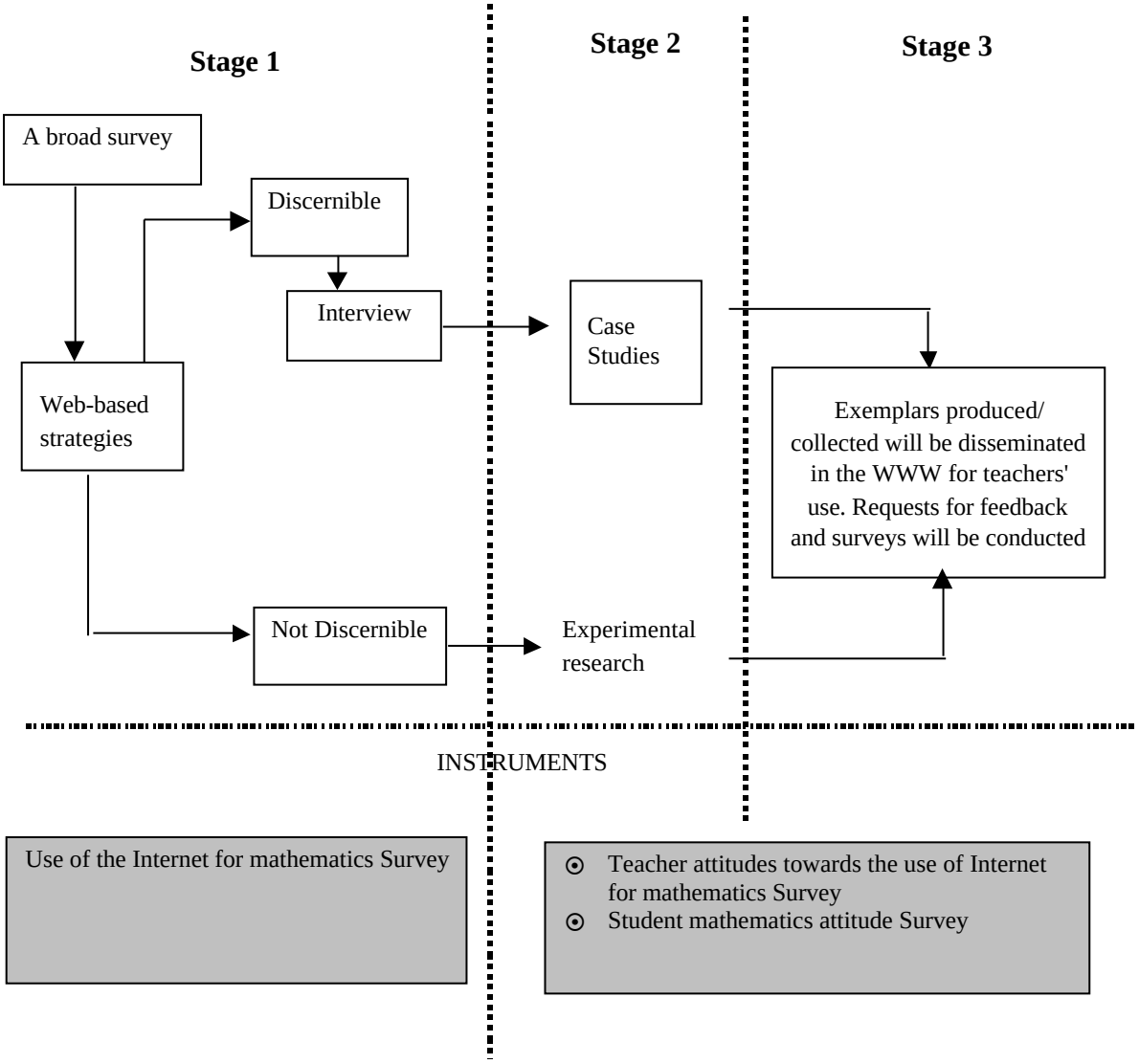


Figure 1. A diagrammatic representation of the methods to be used

It is anticipated that interviews (with teachers who use the Web for mathematics education) will take place to discuss the strategies they have used or are using. Collaboration and observation will begin with interested parties in their respective settings. If there are no discernible strategies, an experimental approach, which incorporates a Web-based unit of work, will be introduced and tested for its effectiveness.

Pre- and Post- attitude surveys and achievement tests will be conducted with samples of Year 10 students whose teachers use the Internet for teaching Mathematics. Year 10 teachers and students have been used because this is the entry level from which many of the concepts of higher mathematics will be introduced. Pre- and Post teacher attitudes' surveys will also be administered.

All surveys are available online in a Web site created for this study. The uniform resource locator (URL) of the site is www.education.unisa.edu.au/elearn/w3mathsed/default.htm.

DISCUSSION

Towards a typology

As teacher use of the World Wide Web for mathematics teaching and learning is relatively new, it is necessary to consider carefully the types or kinds of learning objects. Differing learning objects need to be categorised and typified. The concept of learning objects is used to ensure a link between the utility of the resource in teaching practice and the features and properties of the object. In some cases a learning object may correspond to a web page, in others it may be correspond to more than one page, in still others a single web page may have a number of learning objects useful for mathematics teaching.

It is instructive to consider an example. The following Web page http://www.exploremath.com/activities/Activity_page.cfm?ActivityID=40, entitled "Reflections of a quadratic function" has been classified as interactive and exploratory. The user can explore and compare the graphs of y = f(x), y = -f(x), y = f(-x) , and y = -f(-x) for quadratic functions of the form f(x) =ax_ + bx + c in the applet on the right hand side of the screen. The parameters a, b and c can be changed on the sliders and the user can view the corresponding changes on the graph. Comparisons of the different functions are available by clicking on the required box and a resultant graph will be shown in a different colour. This enables any one of the four graphical functions to be displayed individually or collectively for comparisons and investigations to be carried out. There is also a "calculation/copy" function where values of the different functions mentioned above can be calculated for values of x in step sizes of 1. Cross hairs values of the graphs (necessary for inflection or intersection points) can be obtained by clicking on the "show/hide cross hairs" icon. The graphs produced or the entire screen can be copied on to the clipboard for use in other applications. This page is also linked to a lesson plan, which is available in lab version or lecture version. The lab version is for use in a networked computer classroom whereas the lecture version is for use with a single computer and overhead projector. There is also a corresponding worksheet for student assignment. Clearly interactive engagement with student and discovery learning is supported by such a learning object.

Extensive examination of the some 50 sites suggests a number of groupings of learning objects. Figure 2 shows in graphical form a typology of learning objects on the Web. Two groupings emerge, the first are what might be called "resources" to assist in teaching, and the second comprises "communications" possibilities for mathematics teaching. This diagram shows that apart from having most of the features that are available in stand-alone mathematics software packages, the World Wide Web offers additional features that are peculiar only to it due to its communicative nature.

Interactivity

Interactivity of resources is the key discriminator in their educational usefulness and student activity. The interactive resources can be further categorized as having functionalities which: gives feedback to the user, have exploratory and investigative functionalities, and those, which feature game play.

There are also a comparatively large number of non-interactive sites of mathematical value. These are mainly rich in various types of mathematical information. These consists of articles of historical aspects of mathematics, current mathematics and research issues, archives of topic-categorised word problems, math-related stories, jokes and cartoons and information on free downloads for mathematical use. The other category of non-interactive resources on the Web is the flat 2 or 3-dimensional graphic representations of mathematical objects in areas such as Geometry and Statistics. These are usually colourfully presented and in some cases animated. Most of these graphics are presented in conjunction with explanatory notes and further links.

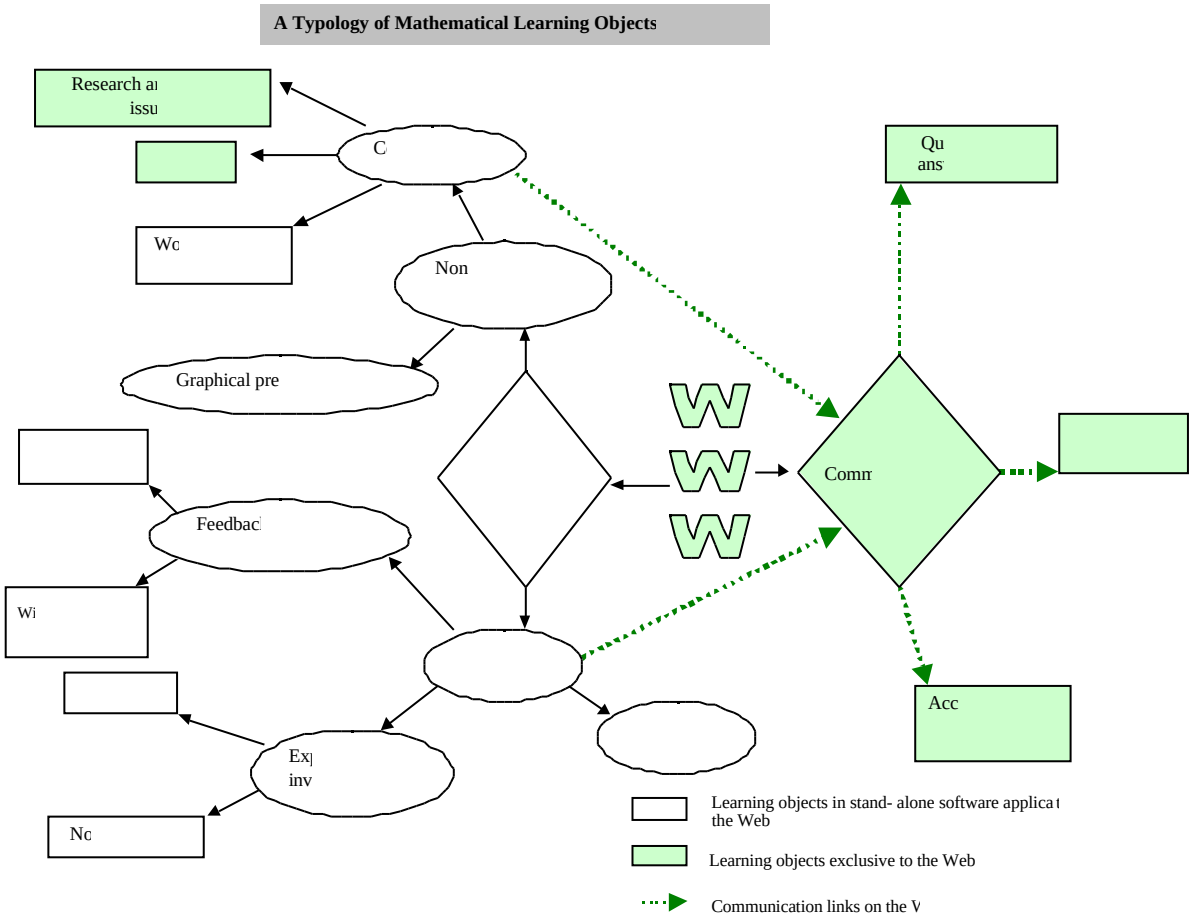


Figure 2. Towards a web typology of mathematical elements

The other grouping features in the communications aspect of the Web. Here the user can participate in online discussions locally or globally; or expect an answer to a mathematics problem from a panel of mathematics educators. These communications are usually made available from the web sites or web pages as links to a forum, a discussion group or Ask an 'expert' (for example, Ask Dr Math).

Some exemplars of these elements can be found in the uniform resource locaters (URLs) presented in Table 1.

Strengths of the World Wide Web

Most software package caters to a particular branch of mathematics; for example, *Geometers' SketchPad* is used for Geometry and *Mathematica* or *Maple* for Algebra. However these are limited by several factors such as licensing procedures and accessibility. On the other hand, the World Wide Web offers a plethora of activities similar in kind to those in stand-alone packages and many of these are public domain. It is a mode of delivery that allows for 'one-stop' accessibility to various forms of mathematical information and activities.

In addition the communications aspect of the Web broadens the possibilities for student learning by allowing the student to extend beyond the limits of the software to gain access to expert opinions, peer discussions and opportunities for obtaining answers to questions.

The World Wide Web is a rapid and relatively cheap means and place to publish educational material. Educational material, such as research and current issues on Mathematics and education, mathematics related jokes and graphics such as cartoons and tried and tested lesson plans, so readily available on the Web, are not correspondingly available in mathematics stand-alone software packages. The vast amount of educational information available on the Web is

something that is not possible within a single CD ROM. The potential also exists for Web based learning environments analogous to WebCT to be built specifically for mathematics teaching.

Table 1. Exemplars of mathematical learning objects on the WWW

Interactive Resources		URL
Feedback systems	With follow-up	http://www.mathcounts.org/GoFigure/Main.taf?function=levels
	Without follow-up	http://cne.gmu.edu/modules/dau/math/dau1_frm.html
		http://cne.gmu.edu/modules/dau/algebra/exponents/exponents_frm.htmlhttp://cne.gmu.edu/modules/dau/algebra/exponents/exponents_frm.html
		http://www.calculus.net/ci2/practice/?tag=9200438920658
Exploratory investigations	Animated	http://www.mathcats.com/explore/xpl/explosion.html
	Non-animated	http://www.exploremath.com/activities/Activity_page.cfm?ActivityID=40
		http://library.thinkquest.org/3288/myojulia.html
		http://www.mathforum.org/sum95/ruth/sketches/algcalc.sketches.html
		http://www.coolmath.com/graphit/index.html
Games		http://www.coolmath4kids.com/games/
		http://www.mathgoodies.com/puzzles/polygon_cw.shtm
Non Interactive Resources		URL
Information-rich	Word problems	http://www.mathforum.org/library/problems/
	Math information	http://www.c3.lanl.gov/mega-math/gloss/knots/knots.html
		http://www.cs.uidaho.edu/~casey931/new_knot/pix_ropes.html
	Research and current issues	Http://www.mathforum.org/social/index.html
		http://www.mathforum.org/mathed/index.html
	Lesson plans	http://www.exploremath.com/lessonplans/
	Freebies for web site	http://kosmoi.com/Science/Mathematics/Graphs/Encyclo/
	Math related-Humour	http://www.csun.edu/~hcmth014/comics.html
	Graphics presentations	http://library.thinkquest.org/3288/mandel.html
		http://www.coolmath.com/polyhedra.htm
Communications		URL
Question and Answer Panels		http://www.mathforum.org/dr.math/
Discussion Forums		http://www.exploremath.com/forum/threads.cfm
Access to Expert Opinions		http://www.mathforum.org/dr.math/

CONCLUSION

The literature shows that there is much research on the use of the Internet and the World Wide Web for teaching and learning in higher institutions of learning as well as for distance education. However little is known about its use in high school mathematics education. This is evidenced by the lack of publications thus far in this area. The intention of this paper has been to whet the appetites of mathematics educators, especially middle and high school teachers or home schooling parents, to seriously consider the usefulness of the Web as a useful resource for extending mathematics teaching and learning. Further study will yield some insight into how the Web has been used or will be used in this area. The typology presented is expected to evolve with time and developments of the World Wide Web.

REFERENCES

Barron, A. E. and K. S. Ivers (1998). *The Internet and Instruction. Activities and Ideas.* Englewood, Colorado, Libraries Unlimited, Inc. and its Division Teacher Ideas Press.

Butler, D. (1998). *Using the Internet- Mathematics*. Cambridge, Pearson Publishing.

Evans, G. (1991). Lesson Cognitive demand and student processing in upper secondary mathematics. *Learning and teaching Cognitive Skills*.

Fuys, D., D. Geddes, et al. (1988). "The van Hiele model of thinking in geometry among adolescents." *Journal for research in mathematics educatio*n.

Hansen, V. P. and M. J. Zweng, Eds. (1984). *Computers in mathematics educa*tion. Reston, Va., National Council of Teachers of Mathematics.

Miller, T. W. and J. M. Miller (2001). "Educational leadership in the new millenium: a vision for 2020." *International Journal of Leadership in Education* **4**(2): 181-189.

Pereira-Mendoza, L. and B. Kaur (1999). "TIMSS-Performance of Singapore Secondary students Part A: Algebra, Geometry and Data representation, analysis and probability." *Journal of Science & Mathematics education in South East Asia.* **22**(2).

Quinlan, L. A. (1997). "Creating a classroom kaleidoscope with the World Wide Web." *Educational Technology* **37**(3): 15-22.

Selwyn, N. (1998). The Internet and schools: An uneasy alliance? [Online] http://computed.coe.wayne.edu/Vol4/selwyn.htm. Accessed 23 May 2001.

Sexton, D. (2001). "Internet enabled modern foreign language learning." *Computer education* **97**(Feb.).

Skomars, N. (1998). *Educating with the Internet : Using Net resources at school and home*. Massachusetts, Charles river Media Inc.

Stager, G. (1996). *Computing and the Internet in Schools: An International Perspective on Development s and Directions.* Victoria:, Incorporated Association of Registered teachers of Victoria.

Williams, M. and K. Price (1999). "Teacher Learning Technology Competencies." *Australian Educational Computing* **14**(2).

Wolfram Research, Inc. (2001) Online *Mathematica*-Based Courses The First of a New Generation of Interactive WebSites. [Online] http://www.wolfram.com/products/mathematica/experience/danishcourse.html. Accessed 10 November, 2001.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

The relationship between values and learning

Bobbie Matthews
Flinders University School of Education

Values are seen as antecedents of behaviour. The study reported has assumed that values have a direct influence on learning behaviour. Values that guide students' lives in their home countries are argued to influence the ways in which learning takes place in other countries. Ethnically Chinese students in Australia are the subjects of the study. Values have been measured on five occasions using the Chinese Value Survey *(CVS) (Chinese Culture Connection, 1987). Approaches to study have also been measured on five occasions using the S*tudy Process Questionnaire *(SPQ) (Biggs, 1987). This paper considers the results of canonical correlation analysis using the four value scales of the CVS and the six SPQ scales. A strong first canonical factor is found in which a weighted combination of values is related to a weighted combination of learning motivations and strategies. The second pair of canonical factors relates low integrity values with a surface learning strategy. The third pair of canonical factors involves low Confucian values and a high deep learning strategy. The second and third canonical factors suggest that for some students their values and approaches to learning are changing during their period of study in Australia. This possible change in values and learning approaches warrants further investigation.*

canonical analysis, life values, learning strategies, Chinese Value Survey (CVS), Study Process Questionnaire (SPQ)

INTRODUCTION

Values are seen as mediators of behaviour. The study reported here assumes that values have a direct influence on learning behaviour. Values that guide students' lives in their home countries will be considered to influence the way in which learning takes place. Feather (1986) points to evidence that value patterns have characterised nations and cultures within nations in history and Biggs (2001) has examined differences in learning approaches of different ethnic groups of students.

Ethnically Chinese students are the subjects of this study. Values have been measured on five occasions using the *Chinese Value Survey* (CVS) (Chinese Culture Connection, 1987) Approaches to study have also been measured on the same five occasions using the S*tudy Process Questionnaire* (SPQ) (Biggs, 1987). The question that arises is do students with different values have different approaches and styles of learning?

This paper considers the results of canonical correlation analysis using the four value scales of the CVS. They are Integrity, Confucian Ethos, Loyalty and Wisdom. The study examines the relationships between the four CVS scales and the six SPQ scales. The SPQ scales are Surface Motivation, Deep Motivation, Achieving Motivation, Surface Strategy, Deep Strategy and Achieving Strategy.

THE CHINESE VALUE SURVEY (CVS)

The CVS was developed by Bond and his colleagues (Chinese Culture Connection, 1987) to measure life values that are peculiarly Chinese or Asian. It uses statements of values; for example, filial piety, the obedience and respect for parents and their financial support and honour and respect for ancestors. Other values measured are loyalty to superiors, knowledge/education and respect for tradition.

The CVS has been claimed by Bond and his colleagues (CCC, 1987) to measure Asian values more accurately than other values questionnaires such as that devised by Schwartz (1992). None of these other values questionnaires has sought to measure accurately values that are distinctly Chinese or Asian in nature. Even measurements taken of Chinese samples by Schwartz (1992) did not 'fit' well with assessments of values of people following a Westernised pattern of thinking.

Table 1 shows the value items of the four CVS scales that have been identified (Matthews, 2000). All the items have adequate positive loadings as estimated using confirmatory factor analysis using LISREL. These loadings are recorded for each item in Table 1.

The Integrity scale measures values such as industry, persistence, kindness and knowledge. The Confucian scale measures tolerance, loyalty to superiors, thrift and respect for tradition. The Loyalty scale measures filial piety, observation of rites and rituals and patriotism. The Wisdom scale measures humbleness, moderation or following the middle way and prudence.

The Study Process Questionnaire (SPQ)

Biggs (1999) describes two distinct groups of learners: those who learn for the sake of knowledge acquisition and those who learn to gain a qualification. The first group learns using a deep approach. They study to learn and are motivated to go beyond the basic requirements for passing. Their learning involves a problem-solving approach and their interest carries them beyond a superficial understanding of what they are studying. These students are actively involved in the learning process and the process involves metacognitive activities that narrow any gaps in knowledge acquisition (Biggs, 1999). Learning in this manner is described by Marton & Säljö (1976a, 1976b, and 1996) in terms of phenomenographic understanding. It is a deep process and precipitates a broad involvement in the learning process itself.

Learning behaviour that is conceptually opposite to deep learning is described as 'surface learning'. It involves only as much as is needed to pass an examination or acquire a qualification. Learners using this approach do not achieve the cognitive levels of deep learning. They tend to be passive and uninvolved in the learning process itself. This is described as extrinsic learning or learning that takes place external to the individual, requiring little personal involvement. It is the diametric opposite of intrinsic or deep learning where the student is actively involved in the actual learning process. Education through deep learning involves conceptual change that goes beyond the basic acquisition of information. Whatever approach is undertaken, it is what the student does that is important. The teacher is only a facilitator, a person who enables learning activities and assists students to achieve desired outcomes (Biggs, 1999). Biggs calls learning designed to achieve the desired teacher and student goals a type of entrapment in a 'web of consistency that optimises the likelihood that students will become engaged in appropriate learning activities' (Biggs, 1999, p.64). He calls this network constructive alignment.

Table 1 Cvs Factor Values And Description (N=573)

	CVS I INTEGRITY AND TOLERANCE (Development of Self)	**_x**		**CVS II Confucian Ethos** (Relationships with Others)	**_x**
2	Industry	0.50	3	Tolerance	0.19
4	Harmony	0.04	6	Loyalty to superiors	0.45
9	Kindness	0.35	16	Benevolent authority	0.32
10	Knowledge	0.34	17	Non-competitiveness	0.45
13	Self-cultivation	0.48	22	Keeping oneself pure	0.44
15	Sense of righteousness	0.47	23	Thrift	0.30
18	Personal steadiness	0.07	33	Contendedness	0.41
19	Resistance to corruption	0.18	34	Being conservative	0.63
21	Sincerity	0.33	35	Protecting your 'face'	0.49
24	Persistence	0.27	37	Chastity in women	0.31
25	Patience	0.68	38	Having few desires	0.56
28	Adaptability	0.45	39	Respect for tradition	0.41
30	Trustworthiness	0.44			
32	Courtesy	0.36			
36	A close friend	0.17			
	=0.89			=0.83	
	CVS III Loyalty to Ideals (Social Responsibility)	**_x**		**CVS IV Tolerance and Moral Discipline (Worldly Wisdom)**	**_x**
1	Filial piety	0.34	5	Humbleness	0.26
7	Observation of rites and rituals	0.78	12	Moderation-following the middle way	0.36
8	Reciprocation of greetings and favours	0.68	14	Ordering of relationships	0.66
11	Solidarity with others	0.43	26	Repayment of good and/or evil	0.52
20	Patriotism	0.50	27	Sense of superiority	0.53
31	Having a sense of shame	0.35	29	Prudence	0.25
			40	Wealth	0.14
	=0.77			**=0.69**	

Learning objectives

Learning is criterion-referenced when the learning objectives are clearly understood by both students and teacher. The objectives may be hierarchical or structural in nature as in the SOLO taxonomy (Biggs & Collis, 1982). Thus levels of the SOLO taxonomy provide a hierarchical framework for learning objectives and the associated teaching and learning activities that may be teacher or self-directed. If teaching and the assessment of learning are aligned at the higher levels of the SOLO taxonomy, students tend to adopt a deep approach both to learning and assessment. This leads to changes in the students that are identified by Biggs (1999) as positive learning outcomes.

Table 2 summarises the approaches to learning in terms of a motivation and a complementary strategy for each approach. The characteristics associated with each approach are also listed. Each approach is divided into affective and cognitive groups that form congruent motive-strategy combinations. The essentials of the three approaches are described as follows: if the goal is to avoid failure, learn key facts and principles verbatim, a

surface approach is involved. If the goal is to satisfy curiosity, a need to understand what was going is required, a deep approach is involved. If the goal is to maximise grades, the student needs to make best use of time; an achieving approach is involved. The SPQ is designed to assess students' intentions and perceptions of learning. (Biggs, 2001).

BACKGROUND TO THE ANALYSIS

Canonical analysis is able to analyse two sets of data simultaneously to see if there are strong and meaningful links between the data. The analysis undertaken in this study has been carried out to investigate the relationships between values and approaches to study.

Table 2 Motive and Strategy in Approaches to Learning

Approach	Motive	Strategy
SA: Surface	**Surface Motivation (SM)** is instrumental: to meet requirements minimally; a balance between working too hard and failing.	**Surface Strategy (SS)** is reproductive: to limit target to bare essentials and reproduce through rote learning.
DA: **Deep**	**Deep Motivation (DM)** is intrinsic: study to actualise interest in what is being learned; to develop competence in academic subjects.	**Deep Strategy (DS)** is meaningful: read widely, inter-relating with previous relevant knowledge.
AA: **Achieving**	**Achieving Motivation (AM)** is based on competition and ego-enhancement: to obtain highest grades, whether or not material is interesting.	**Achieving Strategy (AS)** is based on organising time and working space; to follow up suggestions; behave as a 'model' student.

Following Biggs, J. B. (1987a). and Murray-Harvey (1994).

Canonical correlation analysis has been selected as the statistical procedure for the analysis in this paper because it permits the analysis of two sets of variates simultaneously. It allows the examination and testing of the significance of the relationships between the sets of data. The purpose of the analysis is to investigate whether life values and learning are related to one another. Canonical correlations analysis using the SAS (1985) computer program is the preferred analytical tool as it is able to demonstrate clearly whether significant correlations exist between the two sets of measures.

Canonical variate analysis is a multivariate analytical procedure that was first developed about 1935 by Hotelling but remained largely unused because of the complexity of the computations involved in the analysis. With the advent of electronic calculators and computers, canonical variate analysis emerged as the procedure that has become the general analytic method from which parametric statistical procedures such as analysis of variance, principal components factor analysis and regression have been derived. The term 'variate', first used by Bartlett (1941), refers to observed measures and the term 'variable' to latent constructs that are formed from measured observations (Keeves & Thompson, 1997).

Canonical variate analysis is particularly useful in the study of the relationship between sets of variates. The procedure is able to analyse not only the relationships within a set, but also the relationships between the sets as well as the degree of overlap or redundancy between the sets of variates. One set of variates may also act as a predictor for the other set of variates. Moreover, canonical variate analysis may be used when either or both sets contains variates that are continuous, categorical or mixed (Keeves & Thompson, 1997).

In canonical analysis, the variates on each set are weighted to form the first pair of variables. The canonical correlation between this pair of variables is the highest. The second canonical correlation is the highest that can be found between the X and Y weighted composites that are uncorrelated or orthogonal with the first pair of variables. The significance of each pair of

correlations may be tested using Wilks' lambda and an *F* ratio. Likewise, a third or more pairs of canonical variables may be identified and tested to assess whether they are significantly related.

The canonical correlation coefficient between each pair of canonical variables describes the strength of the relationship between the pairs of latent variables. The square (R_) of this relationship estimates the amount of variance of one latent variable that is predictable from the other latent variable in the pair. The relationship between the two latent variables that form a pair may be viewed as causal and therefore unidirectional, whereby the latent *X* variable forms a causal link with its paired *Y* latent variable. Measures of redundancy are used to examine commonality between blocks of variates in an analysis (Cooley & Lohnes, 1971, 1976; Keeves, 1975, 1986; Keeves & Thompson, 1997; Pedhazur, 1997).

Transformation weights and structure coefficients

There are two types of coefficients that help in the interpretation of canonical variate analysis. These are the transformation weights and structure coefficients. The transformation weights are assigned to the variates that form the latent variable in the linear combination of variates. Structure coefficients are the correlations between the derived variables and the original variates. The structure coefficients are loadings that enable the pairs of variables to be identified and related to their respective variates. The sum of the squared structure coefficients enables the proportion of variance estimated by each factor to be calculated (Keeves & Thompson, 1997).

Transformation weights and structure coefficients assist in attaching meaning to canonical variates. The transformation weights are assigned to the original variates that make up the predictor and criterion sets of variables. The coefficients help to identify the related variables in the predictor and criterion sets and are particularly useful in seeking relationships between the original variables and the derived canonical variables (Tatsuoka, 1973).

Variance

Redundancy measures the proportion of variance of a set of variates that is predictable from its paired canonical latent variable.

RESULTS

Relationships between the latent variables

Table 3. Canonical Analysis: Tests Of Successive Latent Roots

Number of roots	Canonical R	R^2	Approx *F*	df	p <
1	0.40	0.16	8.99	24	0.01
2	0.35	0.11	7.23	15	0.01
3	0.24	0.06	4.78	8	0.01
4	0.10	0.01	1.77	3	N.S.
Overall analysis					
Statistics	**Value**	**F**	**Num *df***	**Den *df***	**p <**
Wilks' Lambda	0.695	8.99	24	1965	0.01
Pillai's Trace	0.340	8.77	24	2264	0.01
Hotelling-Lawley Trace	0.389	9.11	24	2246	0.01

N.S = not significant

The statistical analysis was undertaken using SAS (1985). Table 3 shows the results of the tests of successive latent roots for statistical significance. Three of the four possible roots are found to be significant when p is <0. 01. The fourth root is not significant at the 0.05 level. The table also gives values for Wilks' Lambda, Pillai's Trace and the Hotelling-Lawley Trace. These are all measures of overall statistical significance of the canonical analysis.

In the testing of the successive latent roots, the first pair shows a significant relationship and a moderate degree of correlation (0.40). The second pair of variables is significantly correlated (0.35) and is set orthogonal to the first pair. The third pair is also orthogonal to the other two pairs and the two latent variables are significantly correlated (0.24). The canonical correlation (R) describes the strength of the relationship between the latent variables. The square of this relationship (R_) describes the proportion of variance of one latent variable predictable from the other latent variables in the same set (Keeves & Thompson, 1997)

Relationships between the latent variables and the observed variables

Table 4 records the transformation weights, the structure coefficients and the variance extracted for the canonical variables. The transformation weights are the predictor variables and are marked with the letter 'U'. The transformation weights are the numerical paths that link the predictor variates to the predictor latent variables. The structure coefficients are listed under the criterion variables and are marked with the letter 'V'. The structure coefficients are the path between the latent criterion variables and the criterion measures. The predictor variates have paths connected to the latent predictor variables.

Table 4 Transformation Weights and Factor Structure Coefficients of Canonical Variates

	Transformation weights (U)•			**Structure coefficients (V)†**		
Predictor measures	**U_1**	**U_2**	**U_3**	**V_1**	**V_2**	**V_3**
INTEGRSC	0.22	-1.29	0.22	0.76	-0.64	0.09
CONFURSC	0.18	0.05	-1.25	0.71	0.01	-0.70
LOYALRSC	0.35	0.50	0.32	0.82	0.03	0.19
WISDOMRS	0.48	0.61	0.52	0.87	0.27	0.08
Variance extracted	0.63	0.12	0.14			
Criterion measures	**U_1**	**U_2**	**U_3**	**V_1**	**V_2**	**V_3**
SURMOTRS	0.42	-0.01	0.47	0.61	0.37	0.18
DEEPMOTR	0.20	-0.68	-0.04	0.54	-0.50	0.25
ACHMOTRS	-0.12	0.57	0.61	0.45	0.50	0.42
SURSTRRS	0.30	0.51	-0.73	0.61	0.62	-0.27
DEEPSTRR	0.08	0.09	0.75	0.41	-0.25	0.59
ACHSTRRS	0.62	-0.30	-0.52	0.77	-0.26	-0.07
Variance extracted	0.33	0.19	0.12			
Canonical R	0.40	0.34	0.24			
Canonical R^2	0.16	0.11	0.06			
‡Redundancy	0.05	0.02	0.01			

Transformation weights _ >0.10; †structure coefficients *f* > 0.40 underlined; ‡Redundancy of criteria given predictor/latent variable

The first predictor latent variable is a general values factor. The second predictor latent variable involves the contrasting relationship between wisdom and loyalty with integrity and the third predictor latent variable involves the Confucian factor contrasted with the loyalty and wisdom.

The first criterion latent variable is concerned with general learning style. The second criterion canonical factor involves achieving motivation and surface strategy contrasted with deep motivation. The third latent variable involves achieving motivation and deep strategy.

Path Diagram

Figure 1 shows a path diagram of the canonical correlation analysis following an approach suggested by Keeves, (1975, 1986). Four sets of predictors that comprise the four groups of values from the *Chinese Values Survey* (CVS) give rise to a single latent variable that is labelled 'Values'. This relates to the criterion latent variable called 'Learn'. Learn involves the six criterion variates that are concerned with the learning styles proposed by Biggs, (1987b) for the *Study Process Questionnaire* (SPQ). The path coefficients for Values are the transformation weights for the first predictor factor and are shown in Figure 1. The path coefficients associated with Learn are given by the structure coefficients from the latent variable Learn, with a residual effect that is due to exogenous disturbance (Keeves, 1975).

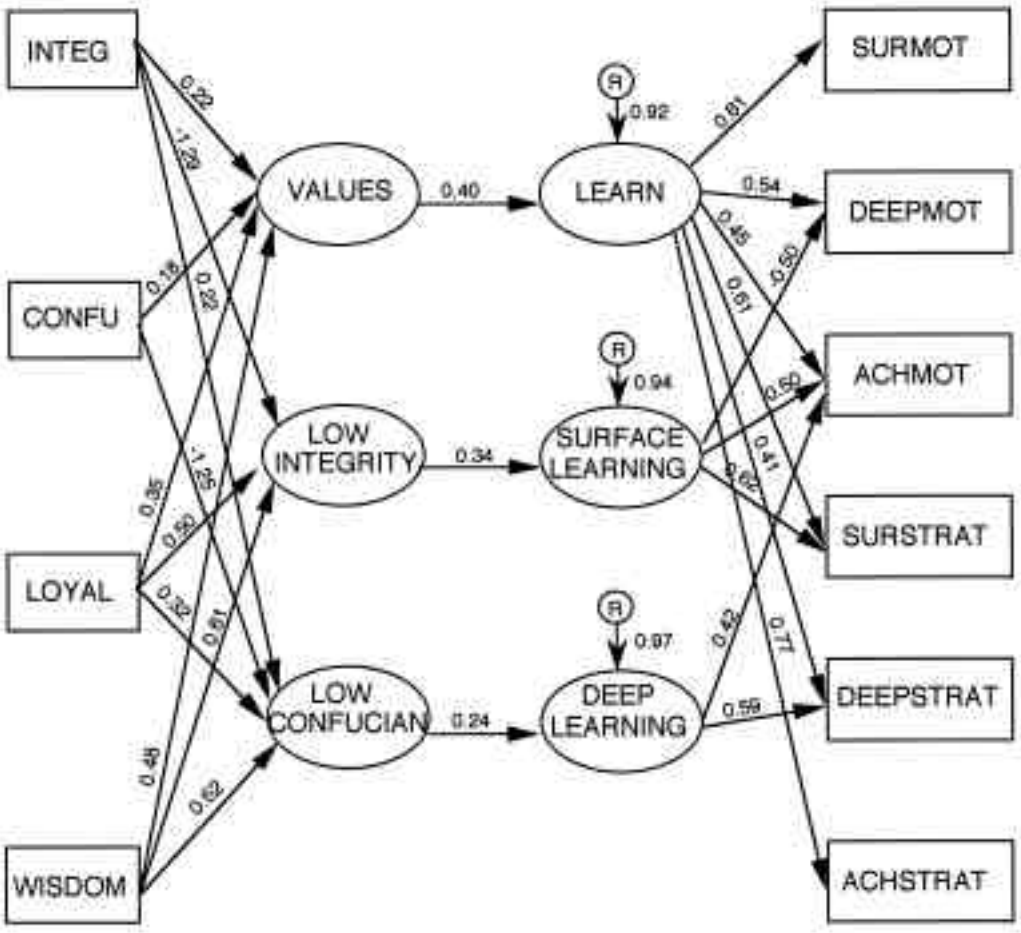


Figure 1 Path diagram showing the correlation between values and learning

The second latent variable, Low Integrity involves three predictor variates, integrity, loyalty and wisdom. The path coefficients for Low Integrity are the transformation weights involved in forming the second predictor variable and are shown in Figure 1.The path coefficients associated with the second criterion factor are given by the structure coefficients that come from the latent variable Surface Learning. Surface Learning involves three variates associated with learning styles, achieving motivation and surface strategy with positive correlations and deep motivation with a negative correlation.

The second factor shows that students who have low integrity values tend to be surface learners. They are low in values such as persistence, industry and sincerity but relatively high in values associated with wisdom and loyalty such as moderation or following the middle way, filial piety, a sense of superiority, having a sense of shame and prudence. These students show a high level of achieving motivation and highly developed surface strategies in their learning. The same students show poorly developed deep learning motivations. They prefer to learn only what is needed to pass examinations or gain qualifications. They do not appear to have a deep commitment to study *per se*.

The third latent variable, Low Confucian, involves predictor variates Confucian ethos and loyalty. This latent variable is formed with the transformation weights given for these variates in Table 3 and the path coefficients for Low Confucian are the transformation weights for the third criterion factor that are shown in Figure 1. The path coefficients associated with the third criterion factor are given by the structure coefficients that come from the latent variable Deep Learning. Deep Learning involves two variates with strong coefficients (values > 0.40), deep strategy and achieving motivation. These are the learning styles associated with low Confucian values.

In contrast, students who are low in Confucian values such as tolerance, non-competitiveness and being conservative, are relatively high in loyalty values such as observation of rites and social rituals and solidarity with others, filial piety and having a sense of shame and on wisdom values such as humbleness, prudence and the ordering of relationships. They show high levels achieving motivation to learn as well as well developed deep learning strategies. These students find it easier to adjust to the ideas and constructs associated with problem-based learning prevalent in the Western university settings in Australia. Their commitment to knowledge and learning for the sake of inquiry into new and different areas of study is likely to enable these students to adapt quickly and easily to university life in Australia.

Redundancy

The results recorded in Table 4 show that the largest contribution to the explained variance for the predictor variables is 63 per cent for the first predictor variable or general values factor, 12 per cent for the second predictor variable and 14 per cent for the third predictor variable. The largest variance for the criterion variables is extracted by the general learning factor of 33 per cent. The second factor accounts for 19 per cent and the third for 12 per cent of the variance.

The redundancies of the criterion variables given the predictors are five, two and one per cent respectively. This indicates that the general values factor explains five percent of the variance of the six learning factors. The Low Integrity and Low Confucian factors only set of variates respectively.

DISCUSSION OF RESULTS

What constitutes good learning

In Australian universities the teacher's role is to engage students in meaningful learning activities that lead to the desired learning outcomes. The teacher should only facilitate the learning process; it is the student who must actively engage in this process for its successful completion. Therefore what the student does is more important than what the teacher does. By implication, there must be active involvement on the student's part for learning to occur. Learner involvement necessitates a problem-based or deep approach to learning rather than a surface approach to maximise optimal engagement by teachers and learners in order to ensure appropriate learning activities and positives outcomes (Biggs, 2001).

This view of learning refers to Biggs' (2001) concepts of what constitutes good learning and what are the optimal conditions for positive outcomes from a Western perspective. This view may require some modification where the Asian learner is concerned. Biggs (1996), Stevenson and Stigler (1992) and others have noted that success and failure in Confucian Heritage Cultures (CHC) is attributed primarily to effort and secondarily to ability. Ability helps but effort is thought to overcome most obstacles to learning. The Chinese proverb, 'failure is the mother of success' encapsulates and underscores CHC thinking. There is an importance of effort in the learning process.

The results obtained from canonical analysis show that lower Confucian values are associated with a deep learning strategy as well as high achieving motivation. In addition lower values associated with integrity and tolerance are related to a surface learning strategy with high achieving motivation but with lower deep motivation.

CONCLUSION

Canonical correlation analysis has shown that two seemingly disparate areas of study, namely, values and approaches to learning, may be analysed and interrelated. The results of the analyses undertaken are statistically significant. Other analyses in this study show clearly that some life values and learning strategies of CHC students studying in Australian universities change over time. This study shows that in general, students with higher values have strong learning strategies. However, students who are low on Confucian values use a deep learning strategy and manifest a strong motivation to achieve and succeed in their studies. Moreover, students who are low in integrity values use a surface learning strategy. They are high in achieving motivation but low on deep motivation. The question awaiting further analysis is whether there is evidence of change in the values and learning strategies of particular groups of students during their periods of study in Australia.

REFERENCES

Biggs, J.B. (1976). Dimensions of study behaviour: another look at ATI. *British journal of educational psychology, 46,* 68-80.

Biggs, J.B. (1987a). *Student approaches to learning and studying. (Research Monograph).* Melbourne: Australian Council for Educational Research.

Biggs, J.B. (1987b). *Study Process Questionnaire Manual.* Melbourne: Australian Council for Educational Research.

Biggs, J.B. (1999a). What the student does for enhanced learning. *Higher education research and development, 18 (1),* 57-75.

Biggs, J.B. (1999b). *Teaching for quality learning at university.* Buckingham: Open University Press.

Biggs, J.B. (2001). Enhancing learning: a matter of style or approach? In Sternberg, R. J. and Zhang, L-F (eds.) *Perspectives on thinking, learning and cognitive styles. (The education psychology series).* Mahwah: Lawrence Erlbaum Associates Inc. (in press).

Biggs, J.B. and Collis, K. (1982). *Evaluating the quality of learning: the SOLO taxonomy.* New York: Academic Press.

Biggs, J.B. and Watkins, D.A. (1996). The Chinese learner in retrospect. In Watkins, D.A. and Biggs, J.B. (eds) 1996. *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp.269-285). Hong Kong: ACER and CERC.

Biggs, J.B. (1996). Learning, schooling and socialization: a Chinese solution to a Western problem. In Lau, S. (ed.) *Growing up the Chinese way* (pp. 147-167). Hong Kong: The Chinese University Press.

Biggs, J.B. (1996). Western misperceptions of the Confucian-heritage learning culture. In Watkins, D.A. and Biggs, J.B. (eds) 1996 (pp.45-67) *The Chinese Learner: Cultural, Psychological and Contextual Influences.* Hong Kong: ACER and CERC.

Chinese Culture Connection. (1987). Chinese values and the search for culture-free dimensions of culture. *Journal of cross-cultural psychology. 18(2),* 143-164.

Cooley, W.W. and Lohnes, P.R. (1971). *Multivariate data analysis.* New York: John Wiley and Sons Inc.

Cooley, W.W. and Lohnes, P.R. (1976). *Evaluation research in education.* New York: Irving Publishers Inc.

Feather, N.T. (1986). Value systems across cultures: Australia and China. *International journal of psychology 21,* 697-715.

Ho, D.Y.F. (1994) Cognitive socialization in Confucian heritage cultures. In Greenfield, P. M. and Cocking, R. R. (eds) *Cross-cultural roots of minority child development* (pp. 285-313). Hillsdale: L. Erlbaum Associates.

Keeves, J.P. (1975). The home, the school and achievement in mathematics and science. *Science education 59 (4)* 439-460.

Keeves, J.P. (1986). Canonical correlation analysis. *International journal of educational research, 10(2),* 164-173.

Keeves, J P. (ed.) (1997). *Educational research methodology and measurement: an international handbook. (2nd edition).* Oxford: Pergamon.

Keeves, J. P. & Thompson, J. D. (1997). Canonical analysis. In Keeves, J.P. (ed.) (1997). *Educational research methodology and measurement: an international handbook. (2nd edition).* Oxford: Pergamon.

Marton *et al*. (1996). Memorising and understanding: the keys to the paradox? In Watkins, D.A. and Biggs, J.B. (eds) 1996. *The Chinese Learner: cultural, psychological and contextual influences.* Hong Kong: ACER and CERC.

Marton, F. & Säljö, R. (1976a). On qualitative differences in learning-I: outcome and process. *British journal of educational psychology, 46,* 4-11.

Marton, F. & Säljö, R. (1976b). On qualitative differences in learning-II: Outcome as a function of the learner's conception of the task. *British journal of educational psychology, 46,* 115-127.

Marton, F. *et al*. (1997). Discontinuities and continuities in the experience of learning: an interview study of high school students in Hong Kong. *Learning and Instruction 7*(1), 21-48.

Murray-Harvey, R. (1994). Learning styles and approaches to learning: distinguishing between concepts and instruments. *British journal of educational psychology, 64,* 373-388.

Pedhazur, E. (1997). *Multiple regression in behavioral research (3rd edition).* Orlando: Harcourt Brace College Publishers.

SAS Institute Inc. (1985). SAS user's guide: basics, version 5 edition. Cary, N. C.: SAS Institute Inc.

Sternberg, R.J. (1997). *Thinking styles.* New York: Cambridge University Press.

Tatsuoka, M. M. (1973). Multivariate analysis in education research. In Kerlinger, F. N. (ed.) *Review of research in education.* Itasca, Illinois: Peacock.

Van de Geer, J.P. (1971). *Introduction to multivariate analysis for the social sciences.* San Francisco: W. H. Freeman and Company.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Indigenous education: Factors affecting students' decisions to continue or withdraw from Tertiary Studies at Flinders University

Doug Morgan
School of Education, Flinders University

This study examined the perceptions of 38 current and former Indigenous students at Flinders University, Adelaide, South Australia, about the factors which influenced their ability to study at university. The study took place at the beginning of second semester 2000. While it was originally intended that students would be surveyed twice circumstances did not allowed this to occur. Preliminary findings suggest that: Indigenous support services are perceived by students as enabling factors in their tertiary studies; younger students are more likely to discuss their tertiary progress with their peers rather than teaching or support staff, unlike older students; female and part-time students believe they would perform better if they had greater domestic support; the lower the Grade Point Average (GPA) the greater the perception of their likely-hood of dropping out and those students who consider dropping out of tertiary studies are more likely to do so.

Indigenous education, Tertiary Studies

BACKGROUND

It is generally agreed that the high attrition rates for Indigenous students in higher education is unacceptable. In general, research on retention rates in higher education tends to focus on the completion and withdrawal of first year students. This is taken as an indicator of overall retention/attrition. There has also been a focus on the implementation of institutional strategies to increase access to higher education by identified equity groups such as Indigenous Australians (DETYA, 1999a). While there has been an overall increase in student numbers over the past ten years due to amalgamations, the falling job market, access and equity policies (such as the Aboriginal and Torres Strait Islander Access Entry Scheme at Flinders University) and the increase in overseas students this has generally not been reflected in university policies. Policies and academic expectations do not reflect factors such as lack of prior preparation or the diverse enrolment patterns of equity groups in higher education (Barrett, 2000). Both these have been found to be factors in student withdrawal, though not all withdrawals are perceived as negative. Some find that the course content does not match their goals, interests or expectations.

McInnis and Jones (1995) found that uncertainty about course choice and dissatisfaction with their university experience has a major impact on students' motivation to study. Price, Hart and Cole (1991) identified other factors. They identified issues such as employment demands, financial concerns, study workloads and fear of academic failure as impacting on student's perceptions of their university experience. These were categorised as internal/external factors by Killen (1994). As would be expected Sharma and Burgess (1994) found an inverse relationship between academic success and attrition. In general, females, mature aged and students who matriculated from government schools were more likely to perform above average and succeed in higher education (Lewis, 1994).

These factors were also present for Indigenous students however, the attrition rates are much higher than those of other categories (DETYA, 1999a). Bourke, Burden and Moore (1996) and Luck (1998) found that Indigenous Support Units in higher education assisted Indigenous students in their studies. Students felt less isolated, less alienated and more at home with the support of Indigenous units. Along with Walker (2000) they also identified personal motivation, dissatisfaction with course of study, lack of family support, financial issues, workload and fear of failure as reasons for Indigenous student's withdrawal from higher education.

This study examined the factors thought responsible for the attrition/persistence of Indigenous students at Flinders University. It was designed in order to provide both academics and support staff with the information to assist intervention with students where appropriate. It aimed to improve the likelihood of Indigenous students continuing tertiary education. Unlike past studies of exit surveys that focussed on the effects of financial, family and personal problems, and issues with low grades and education delivery modes (Bin-Sallik, 1990; Bourke, Burden & Moore, 1996), this study also examined whether or not student identity was a factor in Indigenous students attrition/persistence in university (do students identify with the role of being a student and consider themselves successful) (Slade & Morgan, 1998).

Procedure and Methodology

The study used a modified survey similar to that administered by Dr Salah Kutleileh of the Flinders University Study Skills Centre to NESB and overseas students. It is envisaged that the data collected by Dr Kutleileh can be further compared with the Indigenous and the general student body of Flinders University. This data can then be used to ascertain any similarities between these groups of students and the implications that this will have for the delivery of academic support services (this will be the subject of a later report).

The participants in the study were from two groups - those Indigenous students who were enrolled in first semester but were not in second semester and those who were still undertaking study. The survey was designed gather two types of data. The first part established demographic data such as gender, age, course of study, year level, living arrangements, whether they have a disability, prior education/work history and mode of study (fulltime/part-time). The second part elicited responses to questions exploring the interaction between personal and academic factors and identity as a student on their attrition/academic success. This data was compared with the students self ranked likelihood of withdrawing from studies and whether or not the student had dropped out of university (self reported in the survey or ascertained by the students enrolment status in 2001).

Given that the literature identified the (following) enabling and inhibiting factors we expected to find that:

a) students who consider themselves to be good students were more motivated, achieved higher Grade Point Averages (GPA) and had a lower likely-hood of withdrawal;

b) Indigenous support services would be perceived as enabling tertiary study;

c) family/peer support would be a factor in enabling tertiary study and

d) responding to financial issues would be an inhibiting factor on tertiary study.

All Indigenous students, including those who had not continued on from first semester, at Flinders University were administered a Likert Scale survey (mailed to their home address - see attached) during the first weeks of semester 2, 2000. Students were contacted via e-mail, phone, mail and at student meetings and appraised of the importance of obtaining base data. Those students who do not fill in their surveys were followed up via phone and mail.

Analysis involve descriptive statistics and path analysis to determine the factors effecting student withdrawal and identify the pattern of causal relationships among variables contributing to the decision making process. Data was analysed using SPSS and PLSPATH ver 3.01 (Sellin, 1990) computer programs. All interview data will be subjected to comprehensive content and discourse analysis for more detailed reporting (Hatch, 1992 and Tesch 1994).

Results

Table 1. Gender and Age of Respondents

Gender	Under 21	Over 21	Total
Male	1	9	10
Female	4	23	27
Total	5	32	37

Of these students 26% hold scholarships, 35% are studying part time, 68% have come to tertiary study after working, and 23 % have left employment to come to university (See Table 1).

Analysis of the results showed the following:

1. Students likelihood of dropping out of tertiary studies is significantly higher if it has been considered ($r = .548$ $p = .002$).
2. A good orientation program is perceived as an significant enabling factor in staying at university ($r = -.383$ $p = .044$) as was support from Yunggorendi (Indigenous Student Services) staff ($r = -.519$ $p = .005$).
3. Students rated social life of the university as an significant inhibiting factor on their likelihood of dropping out of university – the better the social life the less the likelihood of dropping out ($r = -.567$ $p = .001$).
4. The higher that students scored goal orientation the less the likelihood of their dropping out ($r = -.457$ $p = .013$).
5. Age was also a factor in the likelihood of student dropping out of tertiary study – those students over 21 years of age were significantly less likely to drop out of university ($r = -.488$ $p = .005$).
6. Students who considered themselves natural scholars had significantly higher GPA ($r = .459$ $p = .024$).
7. The higher a students motivation the their GPA ($r = .417$ $p = .034$).

Discussion of Results

It was not possible to undertake quantitative analysis of differences that may result from gender, age, self esteem, or academic experience due to the limited number students involved in the study. However, the results of the survey indicate that tertiary educators and Indigenous support services play a crucial role in enabling Indigenous students tertiary education. As expected, students who considered themselves to be good students (high self esteem) were found to be more motivated, achieved higher GPA and had a lower likelihood of withdrawal from university study. And of course those students that had lower self esteem were less motivated and more likely to drop out of university study. Those students that indicated that they had considered dropping out of tertiary study were more likely to do so. Female students indicated that negative feedback of tertiary work was more likely to be an inhibiting factor in tertiary study. The role of Indigenous support staff was also found to be an enabling factor for Indigenous students over the age of 21 although female students are more likely than male students to access support staff. Female students are also more likely to discuss their problems with family members than male students were. Indigenous students under the age of 21 tended to prefer advice from their peers when considering dropping out of university. This is problematic as those students who come to tertiary studies straight from school (under 21 years of age) are less motivated than those who come to tertiary studies after

working. It follows then that in order to increase or maintain the academic performance of Indigenous students in tertiary study will require increasing Indigenous students self esteem.

Reflecting the current employment conditions many students indicated that their primary motivation for undertaking tertiary studies was related to increasing their chances of gaining better employment. And financial factors were found to be an issue for students concerns about their ability to continue studies. Part time and female students indicated that lack of domestic support from partners and family were an inhibiting factor in their tertiary education. Students indicated that the social activities of Indigenous support centres were an enabling factor in their studies.

The preliminary indications from this study suggest that if Universities wish to meet equity targets for Indigenous participation in tertiary education then the role of Indigenous support services will be central. It also indicates measures need to be implemented to assist those students who come straight from school to university study. These students are less likely to access support services and are at a greater risk of dropping out of tertiary study. While increasing and maintaining Indigenous participation in tertiary education may entail educators placing greater efforts in building Indigenous student's self esteem, this would appear to be a crucial issue for younger students who are less likely to raise their concerns with Indigenous support staff. Despite a greater commitment to Indigenous access to tertiary education across all university sectors this study indicates that Indigenous support centres are essential in enabling Indigenous students to participate in tertiary studies.

BIBLIOGRAPHY

Barrett, S, (2000) 'The Work/Study Dilemma: a pilot study', in *International Education Journal*, vol. 1, no. 2.)

Bin-Sallik, Mary Ann (1990) *Aboriginal Tertiary Education in Australia*, Aboriginal Studies Key Centre, University of South Australia.

Bourke, C, Burden, Jennifer and Moore, Samantha (1996) *Factors Affecting Performance of Aboriginal and Torres Strait Islander Students at Australian Universities: A Case Study*, Commonwealth of Australia, AGPS.

DETYA (1999) *Selected Higher Education Statistics*

Hatch, E (1992) *Discourse and Language Education* Cambridge University Press Cambridge

Killen, R. (1994) 'Differences between Student's and Lecturers' Perceptions of Factors Influencing Success at University,' *HERD*, vol. 13 (2)

Lewis, D. E. (1994) *The Performance at University of Equity Groups and Students Admitted Via Alternative Modes of Entry*, Canberra: Australian Government Publishing Service,

Luck, Jo 1(998) 'Australian Indigenous students at University: How do we know them?' *Faculty of Informatics & Communication*

McInnis, C. and Jones, R. (1995) *First Year on Campus* , Centre for the Study of Higher Education, Melb. Uni,

Nichols S (1999) *Personal and academic factors affecting students' decisions to withdraw from, or persist at Flinders University*, Research in progress

Price, D; Harte, J. and Cole M. (1991) *Student Progression in Higher Education: A Study of Attrition at Northern Terrritory University*, Canberra: Australian Government Publishing Service,

Slade M D & Morgan D L, (1998) Accommodating philosophical perspective: An imperative for Indigenous higher education, *Proceedings Third Pacific Rim Conference* Auckland NZ.

Sellin, N (1990) *PLSPATH Version 3.01 Program Manual*, Hamburg, W.Germany.

Tesch R (1994) *Qualitative Research Analysis Types and Software Tools* Falmer Press, NY.

Walker, Ros (2000) *Indigenous Performance in Western Australian Universities*, , DETYA

THE FLINDERS UNIVERSITY OF SOUTH AUSTRALIA

Personal and academic factors affecting students' decisions to withdraw, from or persist with study at Flinders University

Name (optional) _____ Student Number ☐☐☐☐☐☐☐

Faculty _____ School/Discipline _____

Please indicate the responses which are most relevant to you by placing a tick (✔) in the appropriate box.

1. Gender ☐ Male ☐ Female

2. Age ☐ 17 – 20 ☐ 21 – 25 ☐ 26 & over

3. Current Living arrangements
   ☐ Living with a partner
   ☐ Living alone
   ☐ Sharing with members of your family
   ☐ Other. Please specify _____

4. Were you born in Australia? ☐ Yes ☐ No
5. If not, please state your place of birth _____
6. How long have you been living in Australia? ☐ Years

Are you an Aboriginal or Torres Strait Islander Australian? ☐ Yes ☐ No

7. Do you hold a Scholarship? ☐ Yes ☐ No
   If yes, please specify _____

8. Do you have a disability? ☐ Yes ☐ No
   If yes, please specify _____

Which of the following corresponds with your background?
- ☐ Non English speaking background
- ☐ International student
- ☐ Local student

Are you currently enrolled at University? ☐ Yes ☐ No ☐ Full time ☐ Part time

11. What course are you undertaking? _____
12. At what stage did you commence
    - ☐ Soon after I completed my undergraduate studies
    - ☐ After working for about five years
    - ☐ After working for more than five years

13. Did you leave full time employment to commence the course you are enrolled in?
    ☐ Yes ☐ No

14. Which, if any, of these statements represent how you see yourself as a student?

Strongly Agree	Agree	Undecided	Disagree	Strongly disagree
5	4	③	2	1

I'm a natural scholar	5	4	3	2	1
I'm an average student, neither brilliant or hopeless	5	4	3	2	1
I could be a good student if I made the effort	5	4	3	2	1
I'm too old to be a student	5	4	3	2	1
I work well with abstract and theoretical concepts	5	4	3	2	1
I could be a good student if my life wasn't so difficult	5	4	3	2	1
I could see myself as a career academic	5	4	3	2	1
I'm bright enough to be a really good student	5	4	3	2	1
I'm a really motivated student	5	4	3	2	1
I don't really fit into university	5	4	3	2	1

15. In this last semester (Semester 1)

Have you considered dropping out of university?	5	4	3	2	1
Have you considered withdrawing from a subject?	5	4	3	2	1
Have you considered reducing load?	5	4	3	2	1

16. When you were thinking about dropping out (withdrawing from a subject/reducing your load) were any of these factors involved?

	Very strongly 5	Strongly 4	Undecided 3	Moderately 2	Not at all 1
<u>Academic Factors</u>:					
negative feedback on work	5	4	3	2	1
low grades	5	4	3	2	1
high workload	5	4	3	2	1
poor quality of teaching	5	4	3	2	1
timetable difficulties	5	4	3	2	1
difficulty of tasks	5	4	3	2	1
style of learning	5	4	3	2	1
<u>Personal Factors</u>:					
loneliness	5	4	3	2	1
stress	5	4	3	2	1
separation from family/friends	5	4	3	2	1
work commitments	5	4	3	2	1
relationship commitments	5	4	3	2	1
a critical event	5	4	3	2	1
health	5	4	3	2	1
disability	5	4	3	2	1

17. When you were thinking about dropping out (withdrawing from a subject/reducing your load) did you discuss your options with any of the following people?

Partner	☐ Yes	☐ No
Friend	☐ Yes	☐ No
Parent	☐ Yes	☐ No
Other family member	☐ Yes	☐ No
Lecturer	☐ Yes	☐ No
Student Counsellor	☐ Yes	☐ No
Study Skills Adviser	☐ Yes	☐ No
Yunggorendi Support Staff	☐ Yes	☐ No

18. How would you rate the likelihood that you will drop out?

Very strongly	Strongly	Undecided	Moderately	Not at all
5	4	3	2	1

19. What changes in your circumstances would make you less likely to leave?

	Strongly Agree 5	Agree 4	Undecided ③	Disagree 2	Strongly disagree 1
increased support from teaching staff	5	4	3	2	1
making friends on campus	5	4	3	2	1
improvement in financial situation	5	4	3	2	1
changes to the timetable	5	4	3	2	1
increased support with domestic/family duties	5	4	3	2	1
getting better grades	5	4	3	2	1
doing your course over a longer timeframe	5	4	3	2	1
a good preparation/orientation program	5	4	3	2	1
more help from Yunggorendi Support Staff	5	4	3	2	1

20. If you decided not to drop out, to what degree have the following factors influenced your decision to stay?

	Very strongly 5	Strongly 4	Undecided 3	Moderately 2	Not at all 1
this qualification will lead to a job	5	4	3	2	1
this qualification will lead to a promotion	5	4	3	2	1
I am getting support from the Yunggorendi	5	4	3	2	1
I am getting support from a study adviser	5	4	3	2	1
I like to succeed at what I do	5	4	3	2	1
The social life at university is good	5	4	3	2	1
I don't want to disappoint my family	5	4	3	2	1
I have promised myself I will finish	5	4	3	2	1
I can't think of anything else to do with my life	5	4	3	2	1
I like the teaching staff	5	4	3	2	1
I am hoping next year will be better	5	4	3	2	1
My partner is encouraging me to stay	5	4	3	2	1
My friends are encouraging me to stay	5	4	3	2	1
My family is encouraging me to stay	5	4	3	2	1

21. If you decided to drop out, would you come back to complete your studies at a later date

☐ Yes ☐ No

22. Are you interested in participating in group discussion(s) that will focus on issues similar to those addressed in this questionnaire? ☐ Yes ☐ No

Please return the questionnaire in the pre-paid envelope provided.
THANK YOU FOR YOUR PARTICIPATION.

What is a 'Good Teacher'?: The Views of Boys

Malcolm Slade
School of Education, Flinders University *malcolm.slade@flinders.edu.au*

The uniform view of 1800 secondary school boys, from 60 South Australian schools balanced across all sectors, is that a 'good teacher' changes everything. In focus group discussions, the views of boys about their declining rates of achievement and retention, quickly turn to teachers. They are profoundly clear about the direct and indirect connections they see between teachers and their own ability to deal with, and to gain control of, other factors influencing both their likelihood of staying at school and their level of achievement. Frequently, the boys ask, with puzzlement and the expectation of an answer that they have been unable to find, why teachers simply can't 'relax', 'loosen up', 'cool it', and 'listen to our view on things'. This paper presents their views about what constitutes a 'good teacher', about ways in which teachers can establish effective educative relationships with boys, and why, with 'good teachers', boys say, 'we want to work' and 'there is not a lot of muckin around' and 'we want to achieve more'.

boys, achievement and retention, good teachers

In response to the declining rates of achievement and retention for boys, the focus of attention, both in research and the implementation of corrective programs, has been on fixing up the boys. The boys themselves, it seems, understand the issues and problems very differently. Although they identify a broad range of factors, central to all of these is that they see themselves to be stuck with an unsuitable learning environment that they cannot change. Based on their experience with *'good teachers'*[1], they see this to be an unnecessary outcome. The fundamental problem, as they see it, is that there are just not enough *'good teachers'*. Putting aside the adult preference to decide whether their views are right or wrong, the aim of this paper is to identify the way in which boys define *'good'* teaching and the impact they believe it has on their level of achievement and their preparedness to stay at school.

The views expressed by the boys are taken from a study involving 1800 Year 9-11 boys from 60 South Australian secondary schools, balanced across all sectors (Slade & Trent 2000; Slade 2000; Trent & Slade 2001; Slade 2001). In groups of 10, the boys took part in one 90 minute open discussion group. Two groups at each school were chosen at random, with one group each from Years 9 and 11. A third group, from Years 9 to 11, was chosen by the school as boys known to be 'at risk'. The boys were asked to draw on their own experiences and to focus their discussion on the issues and problems they believe to be an influence on their level of achievement and their preparedness to remain at school. The factors identified were summarized and made the focus of a second 90 minute discussion with each of the first 60 groups of boys. These boys were asked to critically review the summary, changing it where necessary, to ensure that it accurately presented the views expressed. Toward the end of each session with all subsequent groups, the boys were asked to critically review the ongoing summary in a similar way.

[1] Statements made by the boys are recorded in quotes and in italics.

The views expressed by the boys were clear and uniform across the schools, year levels and levels of achievement. Differences of viewpoint were minor and resulted either from the boys' preferred practice of discussing general issues and problems in terms of local or particular examples, or the varying ways in which individual boys choose to respond. Despite their diversity, for all but a few boys, their general understanding and expression of the issues and problems remained profoundly uniform.

The boys identified a broad range of interconnected factors. Several popularly held views, that the problems start in the primary years, and that the issues and problems are reducible to matters of gender difference, gender equity, peer pressure or literacy and numeracy, have been rejected by the boys as simplistic to the point of being false. Issues about masculinity are conspicuous in their absence.

GOOD TEACHERS MAKE ALL THE DIFFERENCE

> *I reckon that boys are leaving school because of the teachers ... you get pissed off with the teachers and just think 'might as well leave'.* (Year 11)

Despite the broad and complex association of factors, it seems that all but a small number of boys consistently and emphatically see their retention and achievement problems primarily in terms of their relationship with teachers, and what they see to be a proliferation of *'bad teachers'* who don't listen, don't care and who are given too much power. A uniformly repeated view is that a *'good teacher'* changes everything. One good teacher, alone, can make a bad lot tolerable and make achievement, in what is seen to be an otherwise repressive, oppressive environment, seem possible.

> *I think it's hard for guys especially to build a, like, teacher-student friendship or relationship, and that doesn't give a very positive view of school. I mean, like, most might have goals and, like, have a vision for the future, but then because the bridge to that is to do schooling ... a lot of boys tend to pull out because, um, yeah, it just goes back to the teacher-student relationship.* (Year 11)

Not surprisingly, the focus of discussion in all groups either starts out on, or quickly turns to, teachers. All of the boys, to varying degrees, resent what they see as largely ineffective, out of date teaching by people who they think cannot teach, shouldn't be allowed to teach, have lost interest in teaching, and who are unnecessarily, inequitably, inconsistently, and often unsuccessfully, authoritarian.

> *At the moment I'm looking to leave school and get a full time job or something, because I'm sick of all the teachers ... they all harass me all the time ...* (Year 11)

> *I had a teacher last year who didn't really like me all that much ... he didn't like a group of boys in his class ... and if we rocked up after the bell had gone ... he'd shut the door and lock it on you.* (Year 11)

> *Yeah, I had the same problem as well.* (Year 11)

In the main, teachers are seen as structurally empowered, hypocritical bullies who must be defied; often at any cost. Most of the boys are clear about their deliberate intention to display resentment, mostly, and varyingly, in the form of resistance; both as identifiably 'inappropriate' behavior and as a deliberate show of non-involvement in the work. As their achievement goes into decline, the problem compounds and they either *'hang-on'*, knowing they are not performing at their best, or they leave.

> *If you get teachers that are really good, you can chat with them, have a good lesson, then you tend to get more work done. With teachers that are pricks to you, you tend to not like them, not try as hard, retaliate against them.* (Year 11)

> *I'm on probation this year. I wasn't getting anything out of the lessons. I like to make them* [teachers] *feel shithouse. I get a good feeling out of making them really pissed off, cause they do the same to me.* (Year 11)

Although their reasoning shows a lack of experience in some areas, they are profoundly clear about their problems. They are also profoundly clear about the direct and indirect causal connections they see between teachers and their own ability to deal with, and to gain control of, other factors influencing both their likelihood of staying at school and their level of achievement.

> *They don't put much fun and interest into the work.* (Year 9)
>
> *He gives out sheets to people, then sits down and falls asleep.* (Year 9)
>
> *Because we are doing sheets, it doesn't get through to our head ... with Mrs. xxx all we did was nothing but sheets, and that was so boring that everyone in her class just played up ...* (Year 9-11)
>
> *All we do is copy off the board ... no one reads it ... teachers should go through it ...* (Year 9-11)
>
> *When a relief teacher comes in ... our normal teacher has told the relief teacher what we're like and the relief teacher puts it down on us ... without even knowing what we're really like in a different perspective.* (Year 9-11)
>
> *We just mess around in class because we're not learning nothing ... the teachers won't teach us ... it's not interesting ...* (Year 9-11)

Frequently, boys ask, with puzzlement and the expectation of an answer that they have been unable to find, why teachers simply can't *'relax'*, *'loosen up'*, *'cool it'*, and *'listen to our view on things'*.

> *School ... just needs to be a little bit more relaxed ... if you get to a class late or something, or you've got a class you don't really like, or the teacher's had a bad day, so he gets grumpy at you.* (Year 9)

Reflecting on what appears to be an intensely paradoxical predicament, they push the point even further, showing greater anxiety, deep resentment and a despairing need to know why teachers can't do a better job, and why too many of them appear to *'hate their job'*, *'hate the kids*' and care *'only about their pay'*.

> *I hate teachers who hate kids.* (Year 9-11)
>
> *Teachers only care about their pay ... they don't care about teaching you ...* (Year 11)
>
> *Relief teachers don't even know the work ...* (Year 11)
>
> *If you ask a question and you get it wrong, they* [teachers] *go off at you for no reason at all ...* (Year 11)
>
> *They* [teachers] *don't even look like they are going to attempt to help you. They just sit there ...* (Year 11)
>
> *They* [teachers] *don't even explain the work very well. That's probably one of the worst things they do ... they don't explain the work. They just give it to you, and that's the work you have to do for the lesson and they don't explain anything about it ...* (Year 11)
>
> *Our Math's teacher fell asleep in the class ...* (Year 11)

> *Some teachers leave you in the class and just walk out ... they just walk around ... They're just gasbagging with other teachers ...* (Year 11)

> *They* [teachers] *don't care about our futures. They've got their future, so ...* (Year 11)

> *They* [teachers] *tell us that in Year 11, 'You don't have to be here. If you don't want to be here, just go' ...* (Year 11)

> *Teachers teach four or five subjects and they don't know what they're doing ... They're the ones that just hand out the sheets ...* (Year 11)

> *You misunderstand something and they go, like, 'You should have been listening harder', and that's not the reason ...* (Year 9)

> *You can get teachers that give you piles of work, and you're doing heaps of work and it's not really helping you. Then you can get teachers that are a little bit more relaxed, doing stuff that's a bit more interesting, and that's helping you ... and you're not doing as much work. But some people think you have to be writing big essays and all that stuff.* (Year 9)

Despite a wealth of examples to support their negative comments, relationships with teachers are not only talked about in these ways. The boys provide a usefully detailed, consistent account of the things that they believe constitute good as well as bad teaching practice. They do this largely by describing the teachers and coursework that they think either succeeds or doesn't. Perhaps usefully, the profound negativity of their remarks about *'bad teachers'* is consistent with the positive way in which they explain their experiences with *'good teachers'*. A uniformly repeated view is that a good teacher can make a bad lot tolerable and make achievement both desirable and possible.

> *I have a class ... everyone in the class likes the teacher because he's relaxed. He gives us work that's interesting to do, and no one stuffs around in his class because of that.* (Year 9)

Their willingness to identify both negative and positive features of teachers and their educational experience is important, largely because it suggests a relatively balanced viewpoint and a determination to call the circumstances as they find them. Important also, is their preparedness to discuss negative features without the crude assumption that these are necessarily not positive in some way or in some other context.

GOOD TEACHERS CHANGE EVERYTHING

Although the boys list more than 60 defining features of a *'good teacher'*, the emphasis is always placed on the personality of teachers; their ability and willingness to establish relationships of mutual respect and friendship with their students.

It is worth noting that *'slack'* and synthetically *'nice'* or *'groovy'* teachers are not included as *'good teachers'*. In much the same way, *'easy'* or *'bludge'* subjects are not considered *'good subjects'*. In most cases, the subject is as *'good'* as the teacher.

> *We just sit there stuffing around. He doesn't give us stuff to do. I mean, and he does nothin himself. He comes in, yeah, sayin it's boring and 'I'd rather be somewhere doin somethin else', like.*
> *Yeah! But he doesn't give us stuff to do.*
> *But, if you're a teacher and you're students don't perform, you still get paid.* (Year 11)

> *He was a heaps good teacher, but if you get on the wrong side of him, you were stuffed. He stirs you up but doesn't humiliate you.* (Year 11)

Subjects that are regarded as inherently boring and irrelevant are often tolerated, and in some instances actually liked, because the teacher is considered a *'good teacher'*. This is consistent with a more general view that genuine mistakes and imperfections are tolerated from *'good teachers'*. The criteria are decidedly not inflexible and there is no clear 'model' teacher. It seems that teachers are considered *'good teachers'* more because they are seen to be good people in a very general sense, than because they fit rigid criteria in any particular sense.

> *Australian Studies is shit. You just turn off. It's a waste of time. The subject needs to be majorly altered. It's not the teacher* [previously identified as a *'good teacher'*]. (Year 11)

> *And there's nothing better than having a good joke with a teacher, I don't reckon. If you make a teacher laugh, I reckon that's ...* [thumb's-up gesture meaning good]. (Year 11)

> *Mrs. xxx , we joke around with her, and then Mr. zzz comes in, and like, 'You bastards get to work'. She comes in and goes, 'What's goin on. Get to work you little bastards', like, just jokin around.*
> *Yeah! She's alright. You respect a teacher you can talk to. Give respect, get respect. With the strict teachers, you can get into even more shit, and you don't care.* (Year 11)

> *She did suck up to the girls, but she was still a good teacher. We had fun.* (Year 9)

Although the boys express doubts about the value of schoolwork, it remains generally true to say that boys want to achieve. They believe that they can achieve, and that a *'good teacher'* can make this possible.

FEATURES OF 'GOOD TEACHERS': A SELECTION

From the perspective of the boys, the *'good teachers'* are those who, among other things:

- listen to what you have to say;
- treat you like a friend;
- show you respect as a person;
- are able to treat you like an adult, especially in the sense of treating you as they would like to be treated themselves;
- are relaxed and enjoy their day;
- are able to laugh, especially at their own mistakes, instead of trying to conceal them;
- are flexible, but rigorous and consistent;
- explain the work.
- are organized;
- take the trouble to explain their judgements;
- find ways of making the work interesting, and come up with new and interesting things to do;
- respond when boys ask questions in class, and don't make them wait until they have spent time with the students they prefer;
- let you talk and move about in the classroom;
- let you go to the toilet if you need to, and even if you can't prove that you need to;

- don't keep picking on people who have a reputation, pushing them to retaliate by blaming them for whatever goes wrong and niggling them until they do something wrong;
- don't run to the staffroom and swap gossip about kids they don't like;
- don't write out slabs of stuff on the board and expect you all to write it down at the same pace and to learn from it;
- don't humiliate you in front of the class;
- don't pick on or discourage you so that you'll leave school, and don't tell you that you'll never be any good and should leave.
- don't mark you down because of your behavior;
- give you a chance to muck up and learn from it.

With *'good teachers'*, they say, *'we want to work'* and *'there's not a lot of muckin around'*. The problems created by the school environment are much more manageable, and *'we're going to achieve more because we want to achieve more'*.

> *Whatever they do, is what we do. If they're a good teacher and they do better stuff, we do better stuff. If they are a crappy teacher, we do bad stuff.* (Year 9)
>
> *They be good to you, you be good to them ... that's it.* (Year 9-11)
>
> *... they are not completely strict ... no one really talks a lot and there is not a lot of telling off in the class ... Everybody seems to have respect for everyone else and there is not a lot of muckin around.* (Year 11)
>
> *We'll get further with teachers like that ... we're motivated to work if the teacher's relaxed. It makes it fun. We want to work.* (Year 9)
>
> *If the teacher's relaxed we're going to achieve more because we want to achieve more.* (Year 9)
>
> *You feel a lot worse if you haven't done the work for a teacher that you respect than for a teacher that's been bugging you for the last year.* (Year 11)

GOOD TEACHERS ARE TAKING RISKS

From their remarks about *'good'* teaching and from the way in which the boys discuss the necessary features of *'good'* teaching, it is clear that they are identifying teachers who go beyond the policies and pretence of education and its contemporary rhetoric about trust, respect, excellence, valuing difference, or the general demands of thinking in terms of interdependence and relativity.

> *Good teachers are flexible with your behavior. You can joke in class. We drop a couple of words we shouldn't but he doesn't give detentions. He breaks the rules of the school but he doesn't break his own. He's nice to you so you abide by him, we've got respect for him.* (Year 11)

Essentially, they are describing teachers who, professionally and personally, are taking risks by listening to their students, responding, respecting, trusting and valuing their students' views and experiences more than the rules, the policy directives, the legal precedents, their training, their career paths, the reputation of the school, and the views of small but vocal groups of parents. Ironically, the way that the boys characterize these teachers suggests that in many respects they are quietly non-compliant and non-conformist. It also suggests that being seen to be *'good teachers'* may either be the result of them portraying this characteristic(s) or it may

simply be the outcome of their determination to be successful at teaching, as distinct from being institutionally successful teachers. It is also what makes them more valued by students, as positive, culturally relevant role models and often mentors.

The institutionally or professionally non-compliant, non-conformist general characterization, that is most often part of the boys description of *'good teachers'*, might be understood in two ways:

(1) It might be a particular or individual characteristic(s) of these teachers that appeals to most boys and results in them being seen to be *'good teachers'*.

(2) It might also be an outcome, not necessarily one that is wanted, of their commitment to individual, professional and cultural integrity and their determination to be successful at teaching.

Both of these are worthy of further research and offer potentially useful avenues for effective corrective strategies. Nonetheless, it is important to stress that, for two reasons, the first of these should not be confused with the kind of non-compliance or non-conformity that might be the outcome of idiosyncratic, maverick eccentricities, or motivated by a commitment to anti-establishment politics. These are also reasons why it should not be misunderstood as the use of a crude technique to achieve popularity with boys.

- It is a characteristic(s) that appeals to a broad diversity of boys across all schools, all subjects and all levels of achievement.
- It is a characteristic(s) that presents in a broad diversity of teachers, many of whom, in all other senses, might be regarded as quite conservative.

It is also important to note that the second of the above ways of understanding the non-compliant, non-conformist characterization identified by the boys, suggests the presence of a criterion of 'success' or 'successful teaching' that is not only agreeable to both teacher and student, it also makes practical sense out of the rhetoric of education, and may prove to be the benchmark from which all other efforts, to improve the educational experience and its outcomes, should be measured. Paradoxically, it follows from this aspect of their description that teachers who are successful at teaching, are only able to be so by juggling a commitment to two conflicting worlds; one being the real world of the classroom and the other being the hierarchical, bureaucratic and often highly political world of the institution. It is a central claim of the latter, that it exists for the purpose, and with the aim, of making the outcome of the classroom more effective. Nonetheless, it follows from this interpretation of what the boys are saying about *'good teachers'*, that this educative aim can only be achieved by closing the classroom door, ignoring the institutional demands, and getting on with the job of establishing and maintaining trust and mutual respect. Ultimately this is the only effective and reliable basis for the kind of educative relationships that result in real learning and real achievement. *'Good teachers'*, it seems, may be achieving these results at the individual cost of working within this paradoxical dilemma; one that necessarily generates an objective, rational despair and stress.

It is also evident, from the criteria used by the boys to describe *'good teaching'*, that these teachers display a genuine, practical commitment to the ongoing democratization and liberalization of the young; a process that is as much culturally driven as it has been for all other social groups. In doing so, these teachers are effectively offering a resolution to the cultural paradox our students face each day, and to the resulting despair that seems to shape and direct their educational outcomes more than any other factor. In particular, these teachers offer an effective resolution to the nagging paradox that *'school is preparing us for our future, right? But school is way out of date'*. A more threatening dimension of this paradox is that by offering an up to date science, school is both preparing students for a future whilst it offers reasons to believe that the survival of our species is in doubt.

It appears that the central features of *'good teachers'*, are those that offer students reasons to believe that cultural paradox is resolvable. In doing so, they are providing students with sufficient reason to believe in themselves, in their own judgements, and in their future. They also give them sufficient reason to believe in others, in the value of learning and of working toward long term goals. Perhaps more importantly, the boys' experience with these teachers provides reasons to believe that what needs to be done in their lives can be done, and that their confidence in the logic that led to the recognition of paradox, and from there to objective despair, was well founded, rational, but resolvable. Although the impact of *'good teachers'* is that the boys feel *'better'*, they also feel vindicated and genuinely optimistic.

> *For a while, I thought it was just me, that I had problems or somethin. But since I've had Mr. xxx* [a *'good'* teacher] *in Math's, it's all changed ... everythin's better ... even other stuff ... and that was last year. I'd like to get him for everythin. If we had him this year, I reckon I'd do real good.* (Year 11)

From what the boys are saying it is also apparent that *'good teachers'* are making practical sense of the contemporary demands on education, from industry, the community and our increasing involvement in the processes of globalization. They are also making practical sense of the demand that we go beyond the dominant commitment to fragmentation and certainty that continues to direct and restrict the formal educational offering in Australia. In short, these teachers are making practical sense out of the need to think in terms of interconnection and relativity. For several decades, education has managed to make little more than theoretical sense of this need, perhaps because of its appeal to those who are committed more to 'political correctness' than to pragmatic, theoretical and spiritual consistency.

GOOD TEACHERS MIGHT BE MALE OR FEMALE

It is particularly interesting that *'good teachers'* might be male or female. Indeed, the apparent disregard for the gender of teachers is consistent with the boys' view that gender is not a major factor influencing rates of achievement and retention, and that adults concern themselves with gender issues too often, usually in a bid to avoid more fundamental issues. Whether a teacher is male or female is not the fundamental issue, if indeed it is an issue at all. When *'good teachers'* are described, the gender of the individual is coincidental or epiphenomenal.

> *One teacher is excellent. She talks to you like you're a human; you're a person. The rest of the teachers are just like robots. They go to work, get paid, go home. They don't care what you think, what you feel. She will do anything in her power to get you to pass. She gives out her phone number to students who are struggling, so they can ring her and ask her at home. No other teachers do that. They don't care. She's down to earth, got a personality and can take a joke.* (Year 9-11)

Importantly, teachers who express viewpoints, attitudes, judgements or prescriptions that are claimed to be, or seen to be, gender based or gender biased are regarded as *'bad teachers'*, whether the bias is identifiably male or female.

GOOD TEACHERS MIGHT BE OLD OR YOUNG

'Good teachers' are not necessarily young, but it helps. Young teachers are thought to be *'closer to where we are'* and *'enjoy what they are doing'* more than most older teachers. They also *'try harder'* to *'have fun'*, and to make *'the work more interesting'*.

> *But, then again, it seems to be the younger teachers that know what we're goin through at the moment, that seem supportive. Like, Mr. xxx and Mr. xxx. They are really good and they will have a joke and laugh with you. And they won't just say, 'Get lost! - don't want to know you'.* (Year 11)

There are several important aspects to this view:

(1) When the boys talk about young teachers being *'closer'*, this is not explained simply in terms of age in years. Young teachers are more likely to *'treat you like a friend'*, to know about *'the things we're interested in'*, to be *'up to date with computers and other stuff that's important'*, and to understand the kinds of problems that school creates for young people.

More generally, young teachers are thought to be culturally more up to date in themselves, in the sense of being paradigmatically more in tune with the contemporary world. Not surprisingly, teachers who meet the boys' criteria for *'good'* teaching, are often thought of as *'young'* teachers, regardless of their age.

Being young in years is thought to predispose a teacher to be a *'good teacher'* but it is neither a necessary nor a sufficient condition. Not all young teachers are thought to be *'good teachers'*.

Yeah, but some young teachers are living in the past though. (Year 11)

Mr. xxx's not that old and he's probably one of the biggest bastards around.
... he's our year level coordinator and I'm, I'm too scared to even go near him.
He's my Math's teacher and he's always talking about how I ask too many questions.
He's dodgy man! (Year 11)

Nonetheless, teachers who are identified in the boys' discussions as being boring, and who are thought to have ineffective, irrelevant views or methods of teaching, are those who are often said to have had *'too many years on the job'* or who have *'been at it too long and need to go and do something in the real world for a while'*. In contrast, *'good teachers'* are said to be more *'connected'* to the world beyond school, partly because they display signs of this connection in a range of ways, but also because they display less signs of being disconnected.

(2) Importantly, the boys draw a clear distinction between what we might call the 'cultural age' of teachers and their 'age in years'. *'Good teachers'* may be old or young in terms of their age in years but they are necessarily culturally aware, up to date, or connected, and for this reason they are regarded as being *'young'* in terms of their cultural age.

Yeah, Mr. xxx is alright, and he's probably one of the oldest teachers here.
Yeah!
He really is the oldest teacher here.
He's more laid back.
I think it's personal experiences he's been through in his life.
Yeah, exactly.
Yeah, he acts like a twenty year old. He's decent. Yeah. (Year 11)

Teachers who persist in trying to control the lives of the young, whether through approval or disapproval, or by trying to enforce established, traditional or *'preferred'* standards and practices, are talked about as *'ancient'*, *'out of date'* or as *'control freaks'* who are *'out of touch'* or just *'not part of the real world'*.

Primarily, the ideas, attitudes and practices of individual teachers are what distinguishes them as *'good teachers'* or as *'bad teachers'*. In other words, the measure of a *'good teacher'* is the extent to which they are successful at establishing effective, culturally up to date educative relationships with their students; relationships that are based on trust and mutual respect. These are the kind of relationships that reflect the ongoing processes of democratization and liberalization that have transformed the lives and expectations of young people. Given that this is a transformation that is reflected in the lives of the young beyond school far more than within their school experience, the cultural 'connectedness' that boys find in *'good teachers'* is made more pronounced and more highly valued by

being the exception, but not the rule; by being the difference that characterizes *'good teachers'* rather than the norm that characterizes the teachers of *'good schools'*.

In brief, teachers who are more flexible, accepting of change and difference, prepared to value and facilitate diversity, and who display the kind of approval that is not aimed at dominating or controlling the young, are more warmly regarded, irrespective of their age.

> *Mr. xxx is a top bloke. Like, he'll talk to you. Like, he'll talk to us. Like, 'What did you do on the weekend?', 'How ya goin guys' ... Like, we do stuff with him. He's a good bloke.*
> *And you always do well in those subjects.*
> *And then you've got your down right cockheads, like, Mr. xxx. He says stuff like, 'Oh! don't use God's name in vain and all that kinda shit'.*
> *Absolute wankers. Ya don't know how they got to be teachers.*
> *Yeah! Teachers need to be more, like, related to the kids.*
> *Yeah! There's a big generation gap.*
> *There's a lot of older teachers.*
> *A teacher should be more your mate than a teacher, I reckon.*
> *Nah! More like a boss that you can relate to. Like, one you work with and not against kinda thing.*
> *Ya should respect your teacher cause that's someone you want to be like.* (Year 11)

> *Basically all the teachers we've named* [as *'good teachers'*] *are all teachers that socialize with students and don't consider us as students but more as friends.* (Year 11)

Most boys express the view that the majority of their teachers are old in terms of age in years. Although this is not thought to make them necessarily out of date, it does predispose them, and thereby the school environment, to be less in tune with changing attitudes, beliefs and practices, less directed by contemporary challenges, and less focused on preparing for a future that is based on the reality of the present.

> *All those older teachers ... are trying to keep the traditional schooling within, when it is changing - like they are driving a car that way but they are trying to push it back this way. It's stupid.* (Year 11)

(3) Most of all, it is the consistency with which cultural age is displayed, both in terms of viewpoint and practice, that distinguishes *'good teachers'*. Those who are seen to be both old in terms of age in years, young in terms of cultural age, and consistent in the expression of their cultural viewpoint and its application in practice, appear to be the most highly valued and respected.

GOOD TEACHERS ARE JUST GOOD PEOPLE

Despite their ongoing remarks about the urgent need for better teachers and claims that teachers need to be re-trained and brought up to date with the world beyond school, the boys quite uniformly declare that *'ya can't train good teachers'*, because it is mainly a matter of *'their personality'*.

> *If the teacher doesn't see it that way, well, there's a low mark.*
> *Yeah, exactly, and I think it's goin to be damned hard for the government to change the perception of the teachers, like, upon us.*
> *It's not going to happen quickly.*
> *Nah, exactly, it's not going to be okay. Ya can't just go, here's a lot of money, let's do this to make it change. Ya can't do it. Ya can't, like, open their head and say, 'Okay, you've got to do this now'.*
> *Well, you could, but it probably wouldn't shut again.* (Year 11)

It's just their personality. Whatever kinda person they are, really. (Year 9)

Some teachers just shouldn't teach. It's just not their thing. They don't have the right personality, or somethin. (Year 11)

Basically, *'good teachers'* are good people. Good people are easy to be with, work with, work for and to even work hard for; they offer no obstacle to mutual trust and respect, they are consistent in the full sense of the term and openly seek to be so, and they offer no reason to doubt that they genuinely listen and genuinely care.

It's his character. He's just, like, easy to do stuff with. (Year 9-11)

The emphasis in teacher training, from the viewpoint of boys, should be more upon selecting people for teaching who have the right kind of personalities or characters, than upon academic training aimed at credentials and accreditation.

You really need to be a people person to be a good teacher, and like, uni doesn't really teach you that. (Year 11)

This suggests that the boys have no confidence, either in the capacity of *'bad teachers'* to change, or in the tertiary education system to produce more *'good teachers'*. Indeed, from their direct comments, the boys are not optimistic about the likelihood of changing people in or through education because education is more focused on rewarding conformity and compliance, than it is about learning to be *'good'* people. This is particularly true of those who they believe are academically qualified but either just not suited to teaching, or whose personalities are more suited to the kind of teaching that no longer works and can no longer be considered effective teaching.

A more optimistic interpretation of the boys' view might be that they are focusing on particular personality traits and the characteristics displayed by teachers working under particular conditions. The view that these teachers may have more suitable personality traits or characteristics that are either suppressed, not publicly displayed or not professionally applied, doesn't feature strongly in boys' discussions. It appears that most of the boys are thinking that what they see in teachers is all that there is or might be in the personalities of these people.

How they are as a person is how they teach. (Year 11)

The boys seem not to be critically aware that the premise 'How they are as a person is how they teach', does not necessarily support the conclusion that 'How they teach, is how they are as a person'. Although this conclusion may be true, it may also be false. Furthermore, on several occasions some of the boys talked about having had experiences with particular teachers that might have served to challenge their reasoning. One teacher, for example, who had been identified as a *'bad teacher'* was said to have been *'a lot different when we were on camp'*. Of another, it was said that *'on tour last year he was really good, like, nothin like he is at school'*. Nonetheless, these experiences appear to be rare and not sufficient to raise doubts about the general view that the teacher is the full extent of the person. Of course, had these experiences not been separated from the school environment, and had the school environment not been one that is thought of as being detached from the world beyond, their impact may have been greater or less easily dismissed.

It is worth noting that embedded in the boys' reasoning is an indication that most of them do entertain a more positive, optimistic belief that the personality displayed by *'bad teachers'* might actually and easily be changed. The belief is based on a general and strongly held conviction that pragmatic reasoning is universally persuasive and effective. In this case it informs the view that, ultimately, everyone involved with schooling wants a better deal. The reasoning they use goes roughly like this:

- The things that make someone a good teacher are obvious.

- *'Good teachers'* get better results from students and enjoy their job more.
- The outcome is that school is a better experience and results in better outcomes for everyone.
- Hence, most *'bad teachers'* will want to be *'good teachers'*.

We can see this kind of reasoning embedded in a view expressed by a Year 9 boy. He had been involved in a school stage production, and missed a lesson. He found the teacher during a lesson break and, although he was given a brief outline of what was required for homework, he remained unclear, and found it difficult to do the work correctly.

> *I did the wrong thing for homework. I misunderstood what he said. So, I told him I did the wrong thing and I didn't really want to read it out. So then he yelled at me and wanted me to read it out, but I didn't want to read it out because I was embarrassed about it. So then he told me to get outside ... It doesn't help to go outside, you're just missing more of the lesson.*
> *You shouldn't get into trouble for things like that ... they should just read it after class. If you miss some work you should be able to catch it up at home or something, you shouldn't get blasted at in class.*
> *I have a class ... everyone in the class likes the teacher because he's relaxed. He gives us work that's interesting to do, and no one stuffs around in his class because of that.* (Year 9)

In discussions with staff at participating schools it has often been argued that teachers are prevented from being the kinds of people that the boys are talking about as *'good teachers'*; that they are forced by institutional and professional constraints to be policy directed, boring, authoritarian *'robots'*. Many teachers said that they shared the boys' views on most things, and that they thought themselves to have been poorly supported as teachers, and inadequately trained from the outset; that their tertiary training and subsequent professional development was, to use the words chosen by the boys to describe most aspects of school life, *'boring'*, *'repetitive'* and *'irrelevant'*.

The boys, on occasions, argue that there is not enough government support for schools, that teachers don't get enough free time to help students, that there are too many students in classrooms, and that teachers are under pressure to cover a set amount of work in a set time, whether or not all students can keep up.

> *If the teachers had less people to teach, like, if they were only teaching, like, two different classes or something, they wouldn't be going, 'Oh, I can't do it this session, cause I've gotta do this for this other lesson'.*
> *And you're having a lesson and your teacher's marking different classes' stuff. Like, thanks for teachin me.* (Year 11)

> *Our teacher tries to rush it through, to keep up.*
> *It's cause they have to keep with the time schedule, otherwise they would be able to go through it more thoroughly and you would be able to learn it better.* (Year 11)

From their discussions, I find no reason not to conclude that all boys would support this view, in principle. They have also argued that what prevail as *'community standards'* in schools are too often those of a vocal minority of parents and influential or empowered staff who bully the rest into complying. Together, these arguments suggest that the boys would support the view that more teachers would be *'good teachers'* if the conditions in schools made this possible.

Nonetheless, the view expressed uniformly throughout the boys' discussions, is that if some teachers can manage to be *'good teachers'* under the prevailing conditions, all teachers can, and those who can't should leave. At some stage in all of their discussions about teachers, the boys return to the same argument, based on their experience with *'good teachers'*.

THE WAY TO GO IS TO LET STUDENTS CHOOSE

The solution favored by many boys, to the difficulties of providing appropriately trained teachers, and of motivating existing teachers to meet the needs of the current students, is simply to give students the freedom to choose their teachers and their subjects. In short, to allow the logic of the market place to decide who should teach and who should not.

Rightly or wrongly, the boys show great confidence in market place logic. They see themselves as consumers and they believe that the market place has the capacity to provide what is best for consumers through the forces of supply and demand. They claim that the adult world has double standards in this respect. They say that they are constantly told by adults, especially teachers, that they need to *'shape up'* if they want to get a job. They are also told that their poor performance at school is the result of a poor effort on their part, that they have nobody else to blame but themselves, and that the world is justifiably competitive and unrelenting in its preference for the best. Although the boys express a strong dislike for the inconsistent way in which many teachers are said to use this argument, they offer no challenge to the argument itself. However, they do want to know why the same argument doesn't apply to the provision of relevant courses, suitable teachers and appropriate school environments.

Many boys in the private school sector are very mindful that their parents pay for their education costs. They say that they get regular reminders of this fact, by both parents and teachers. They repeatedly express the desire to speak for themselves, and to choose as they would do, if they were empowered consumers or clients in any other market place exchange. Some of these boys are of the view that their parents are increasingly inclined to accept their judgement and to support their choices. Indeed, the most satisfied group of boys in this study were at a private college specializing in Year 11 and 12 academic courses. All of these boys had come from another school. Most of them had come from other private schools, seeking better teaching and a better school environment. In all cases, *'better'* was determined pragmatically in terms of finding the conditions under which they would do their best. In all cases, this translated into getting 'what they were paying for' and in the main, this meant *'good teachers'*.

> *I'm only in Year 9 now, but I'm waiting to go to xxx* [the senior college]. *My sister's already there, and she says it's really good. I'd like to go there now, but I have to wait until Year 11. That's nearly two years.*
> *Why can't this school* [his current school] *do what they're doing? If they did, I'd stay here'.* (Year 9)

Although most boys can see ways of changing the curriculum and the school culture, these remain contingent on having more *'good teachers'*. If a necessary condition of getting enough *'good teachers'* is to *'get rid of the bad teachers'*, they see no reason why this should not be done and they see themselves, along with the girls, as well qualified to know who should stay and who should go.

> *There are definitely good teachers and bad teachers. If we could get rid of the bad teachers we'd know who to get rid of.* (Year 9)

In their discussions about *'good teachers'*, the boys were asked how many *'good teachers'* they thought they had experienced in secondary school. At the majority of schools, the number remained around 10% to 20%. Although this is a percentage of the teachers they had experienced, which is less than the total number in the school, it does represent the percentage of their collective experience. At some schools the percentage was higher but rarely above 30%.

> *Eighty percent of the teachers at this school are absolute shockers. They've got no idea ...* (Year 11)
> *Teachers are in it for themselves, basically.*

> *... the thing is that there is only one or two teachers that are willing to help you do better. But it seems to be the majority of teachers that don't give a hell. Like, just, like, go away. I don't want to know about it.*
> *And when they do help you, they've got, like, you can tell. They help you but, like, you can tell by the smirk or the attitude that they really don't want to be there.* (Year 11)

NOT GENUINELY LISTENING, IS NOT LISTENING AT ALL

Most of all, the boys have uniformly expressed the view that the adult world, which in their school experience amounts to teachers, does not genuinely listen to their views and is not genuinely interested in their concerns. They use several reasons to support this claim, but two are of particular relevance:

- *'they*[adults/teachers] *usually just say that we're wrong or we don't understand'.*
- *'they* [adults/teachers] *do nothing about it'* that is *'even when they say they're listening, nothing changes - so they can't be'.*

Given the imperatives of our time and our culture, and the pressing need, locally and globally, to think and act in terms of interconnection and relativity, in and through education, it would seem that the boys are drawing attention to the central aspect of a much broader problem, namely, that we are failing to listen to ourselves, individually, as a community, as a culture and as a species.

In the main, current strategies emphasize the importance of matters like curriculum, assessment, credentials, policies, programs, guidelines, career paths and accountability, focusing largely on fixing up the boys and satisfying a despair driven adult need for control and certainty in a world that is fundamentally understood in terms of fragmentation. The boys, on the other hand, have emphasized the importance of people and their cultural age; their personalities, attitudes, ideas and values, the way that these influence their practice, and the extent to which they display a commitment to pragmatic, theoretical and spiritual consistency. In other words, instead of focusing on the institution, they focus on the individual; instead of focusing on status and the exercise of power and correctness, they are focusing on the extent to which integrity, in all of its individual, professional and cultural senses and dimensions, is retained - wherever, whenever and at whatever the cost.

REFERENCES

Slade, M. & Trent, F. (2000) What the boys are saying: An examination of the views of boys about declining rates of achievement and retention. *International Education Journal*, 1 (3), 201-229. [Online] http://www.flinders.edu.au/education/iej [2000, December]

Slade, M. (2001) Listening to the Boys. *The Boys in Schools Bulletin*, 4 (1), 10-18. [Online] http://www.newcastle.edu.au/department/fac/boys/ [2002, January]

Slade, M. (2001) *Listening to the Boys*. Unpublished doctoral thesis, Flinders University, Adelaide.

Trent, F. & Slade, M. (2001) *Declining Rates of Achievement and Retention: The Perceptions of Adolescent Males*. Report to the Department of Education, Training and Youth Affairs, Evaluations and Investigations Programme, Higher Education Division, Commonwealth of Australia, Canberra. [Online] http://www.detya.gov.au/highered/pubgen/pubsalph.htm [2001, August]

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Validation of the Aesthetic Judgement Scale through an investigation of interrelationships with external predictors

Joanna Touloumtzoglou
Flinders University, School of Education *ctsopela@bigpond.com*

A sample of 478 secondary school students were assessed with an Aesthetic Judgement Scale (AJS) for the purpose of measuring performance in art production. In order to establish the validity of the AJS measurement and the scores generated, the meaningfulness and strength of these measures as well as their relationships to other external standards required examination. A causal model of secondary student performance in art production was therefore advanced to investigate relationships between multiple predictors and the criterion variable of performance in art production (AJS). The model included structural, psychometric, and attitudinal variables hypothesised to contribute to individual differences in performance in visual art production. A battery of psychometric tests was employed to test the hypothesised model that measured performance on the types of constructs indicated. The model was subsequently tested with partial least squares path analysis (PLS) to confirm the theory and the hypothesised relationships. Separate analyses were undertaken for each grade level, in order to examine the constancy in the characteristics measured, and the uniformity of the assignment of measures across grade levels.

Partial least squares path analysis, art production, latent variable, manifest variable, direct effect, indirect effect

INTRODUCTION

Educational researchers support the view that outcomes of art education and complex behaviours involved in visual expression are measurable constructs, while acknowledging the necessity for improved measurement (Clark & Zimmerman 1992; Hardiman, Liu & Zernich, 1992; Armstrong 1994; Beattie 1997). There are many who oppose the use of a single criterion in assessing artistic potential, while unanimously supporting multiple predictor assessment procedures (Renzulli & Smith 1977; Chetelat 1981; Renzulli, Reis, & Smith 1981; Renzulli 1982; Saunders 1982; Clark & Zimmerman 1984; Ghallagher 1985; Wenner 1985; Boston 1987; Ellam, Goodwin, & Dought 1988; Hurwitz & Day 1991). The predictive value of measures obtained with intelligence tests, creativity tests as well as student background characteristics, when screening for artistic potential, has been corroborated by many researchers (Burkhart 1962; Beittel & Burkhart 1963; Csikszentmihalyi & Getzels 1971; Clark & Zimmerman 1988; Scott 1988).

Assessment of performance in visual art production needs to address the diverse abilities contributing to artistic potential. Jellen and Verduin (1986) argued that the multifaceted nature of artistic potential should be defined as a mental construct that is comprised of cognitive, affective and conative abilities. Barron (1972, 1974, 1977, 1982) has systematically researched the types of tests that are the best predictors of artistic ability. Positive relationships have been identified between ratings of art portfolios with performance on measures of perception accuracy, analytic powers, student attitudes and aesthetic judgement (Barron 1982). The basic

design employed for his predictive studies involved the use of portfolios of art work rated for artistic potential, performance measures on a battery of psychological tests, faculty ratings, grade-point ratings, honours and scholarship support. All the information collected was used to construct a composite criterion for prediction.

Overall, there is substantial evidence that identifies a relationship between the domains of visual perception, figural creativity, visual-preference and attitudinal measures, as well as home background characteristics and the processes of art production.

Central to this study is the validation of the Aesthetic Judgement Scale (AJS), an instrument developed for measuring secondary school students' art production, for assigning quantitative specifications both singly and in combination to corresponding magnitudes of art work. The construct validity of the AJS was perceived as being the extent to which the instrument measured the constructs it purported to measure (Kaplan 1964). Validity therefore, could be established from the meaningfulness and strength of the relationships between the AJS scores and other external standards. For this purpose measures of school achievement, psychometric, attitudinal and structural (home background) variables were obtained and included in a multi-component path model for assessing secondary school students' art production incorporating measurement obtained on the AJS.

DATA COLLECTION

Data were collected from a secondary school with an international character and a wide range of differing abilities within the student population. The school provided specialised programs catering for students of high intellectual potential (SHIP) with a central focus on enhancing divergent and creative thinking, particularly relevant to this study. The total number of students in the final sample employed was 478, including 238 males and 240 females ranging from Grades 8 to 12. Sources of data involved measures of: (a) school achievement variables; (b) psychometric variables; (c) attitudinal variables; and (d) home background variables.

School achievement variables included measures of performance in art production obtained on the Aesthetic Judgement Scale (AJS) as well as the grade point average of students' scholastic achievement (GPA). The Aesthetic Judgement Scale was developed for assessing secondary school students' artwork across the five grade levels (Grades 8–12). Succinctly worded level descriptions were provided for assessing student level of performance and for allowing comparisons to be made between past and present performance.

The AJS is a polychotomous rating instrument where markers classify student art production against a five-point scale of performance on 25 items. In this way, an estimate of a student's level on the achievement continuum is produced that confers partial credit for partial success. Hence, for a rating of Extremely Poor students are assigned a code of 0; Poor corresponds to a code of 1; Fair corresponds to a code of 2; Good corresponds to a code of 3; and Extremely Good receives a code of 4. The inclusion of a number of categories as well as their more meaningful use increases the precision in the ratings.

The AJS was employed as a culminating assessment instrument on two occasions during the visual arts course, since on-going assessment of student artworks was not feasible. A majority of students in the sample undertook the visual arts course over one semester only, necessitating measures to be collected at the end of each term, with the added value of examining relative growth between two occasions. A completed piece of art work was assessed with accompanying back-up material, demonstrating the process of idea and imagery development. The art work was selected by the art teacher as being the most mature and coherent piece of work that best represented the student's level of performance. An examination of all accompanying authentic material including journals, notes, sketches, and evidence of research, assisted in the assessment of the final piece of art work.

Students' GPA involved an aggregate of marks on each subject undertaken during the school year that was calculated as a measure of general scholastic achievement. Since marks were awarded in the form of alphabetical symbols, their conversion into numbers was required before estimating a grade point average (GPA). Therefore numerical values were assigned to each alphabetical symbol (A = 4, B = 3, C = 2, D = 1, F = 0), then the aggregate was estimated over the number of subjects undertaken and multiplied by 100. Students having completed subjects for the International Baccalaureate (IB) program where marked on a separate assessment scheme with scores ranging from seven to nil (where A =7-6, B =5, C =4, D =3 and F =2-1) which were also aggregated using the same procedure.

For obtaining data for the psychometric variables, a battery of psychometric tests were administered including: (a) the *Torrance Test of Creative Thinking* (Torrance 1974) for assessing figural creativity; (b) the *Visual Aesthetic Sensitivity Test* (Götz 1985) that provided measures on students' good taste factor; and (c) the *Hidden Figures Test*, *Card Rotation Test*, *Surface Development Test*, *Gestalt Completion Test*, and *Identical Pictures Test* (French, Ekstrom & Price 1963) for the assessment of visual spatial abilities.

Attitudinal variables involved the collection of data from the administration of the Attitude Scale toward the Visual Arts School Subject[1] (ASVASS) and the Visual Arts Attitude Scale[2] (VAAS) respectively, as well as aggregates of teacher ratings of students' scholastic deportment assigned for each school subject. The calculation of these aggregates followed a similar procedure to that of the GPA stated earlier. Four ratings were used by teachers to mark students on the basis of attitudes, effort and deportment, each being assigned a numerical value: *Excellent* = 3; *Good* = 2; *Satisfactory* = 1; and *Unsatisfactory* = 0. These ratings were also averaged over the number of subjects undertaken and multiplied by 100.

Ultimately, information relating to students' background was collected through the administration of a Home-Background questionnaire that provided information on structural variables and home environment characteristics.

METHOD OF ANALYSIS

Partial least squares path analysis (PLS) (Wold 1982) was used in this study so as to examine how useful and meaningful AJS[3] measures were in assessing student learning and development in art production, once direct and mediated effects of other intrinsically valid measures associated with the criterion were taken into account. This was achieved through the identification of how variables mediated effects of other variables, and how variables were interrelated. Moreover, in order to examine the constancy in the facts to be measured and the uniformity in the assignment of measures across grade levels, separate analyses were undertaken for each grade level. This was brought about through: (a) replication; (b)

[1] An adaptation of a scale measuring attitudes toward any school subject (Silance & Remmers 1934).

[2] An adaptation of a scale for the measurement of attitude toward the aesthetic value (Cohen1941).

[3] Results of confirmatory factor analysis identified an underlying higher-order *artistic aptitude* factor indirectly measured by the observed variables and affecting three more specific factors, namely: (a) *elements of art*; (b) *formal properties*; and (c) *symbolic level.* An additional four observed variables were also shown to reflect an alternative form of the *artistic aptitude* factor and were represented by an exogenous model labelled the *composite* factor. All in all, the general underlying dimension (artistic aptitude) implied that it was sufficient to calculate a total score for the measurement of student performance on the AJS.
The item-response data were Rasch-scaled to establish further the dimensionality of the AJS and to allow for an estimation of performance on a calibrated interval scale that was independent of the difficulty distribution of the items sampled, and of the ability distribution of the persons sampled.

examination of change or development across grades; and (c) investigation of how grade subgroups operated when self-selection effects were controlled[4].

PLS path analysis is a multivariate technique that allows modelling of relationships among multiple predictors and the criterion variable, as well as the construction of unobserved latent variables (LVs) (Chin & Newsted 1999). This technique provides for the combination of manifest or observed variables (MVs) that may correspond to actual items in tests and questionnaires employed for data collection, into latent variables (LVs) in order to reduce the effects of measurement error, and to employ a more general construct instead of a specific observed variable reducing the number of cases necessary to obtain stable estimates (Lietz 1996).

PLS was employed because it is distribution-free, allowing all variables for inclusion in the proposed model to be entered regardless of their distributional characteristics. From a necessity to estimate relative change in the AJS measures between assessment occasions, variables representing student scores at each occasion were likely to differ. Consequently, an examination of the same model required replicating analyses for two systems on the two separate occasions T1 and T2 per grade level.

The computer program PLSPATH version 3.01 (Sellin 1990) was used in this study for analyses of the causal models of student-level factors influencing performance in art production. The advantages of this computer program comprise: flexibility of use with raw data; quick computation of direct and indirect effects; and a capability of handling many variables simultaneously. There are, however, approximate limits, of the case to variable ratio. A sample size requirement is either ten cases per formative indicators in a block with the largest number of MVs, or ten cases per independent LV influencing the dependent LV with the largest number of LVs (Chin & Newsted 1999). For this reason, in estimating the LVs hypothesised in this study, successive models were employed where the required cases to variable ratio (10:1) was satisfied. However, in the analyses of data undertaken separately for each grade level, this ratio constraint did not permit estimation of parameters to be undertaken with Grade 12 data because of the small sample size (N = 27). Consequently, the results reported in this paper were derived from the analyses of data from Grades 8 to 11 only.

RESULTS

Path analyses of data by occasion and grade level

The hypothesised path model

A concise summary of the MVs and LVs employed in this study, as grouped under their respective dimensions, is given in Table 1. The table is sub-divided into seven dimensions and the criterion variable. These comprise the: (a) structural dimension, which includes parents' occupation status, education level or other characteristics such as ethnic background, gender or age; (b) attitudinal dimension of the home, which involves parents' attitudes and expectations; (c) process dimension of the home, which encompasses what parents do that influences students' performance; (d) attitudinal dimension at the student level, which involves students' attitudes; (e) process dimension at the student level, which includes what students do that affects their performance; (f) domain-relevant abilities, which concern students' visual spatial abilities, and their performance on a visual aesthetic sensitivity test and a figural creativity test; and (g) the scholastic achievement dimension, which is reflected by students' grade point average (GPA).

[4] Self-selection bias is noted where students in Grades 9 to 12 are allowed to choose whether to undertake a visual art subject during the school year, contrasted by students in Grade 8, who are obligated to do so.

The path models in Figures 1 through 8, contain the criterion variables AJS1 (AJS scores at T1) and AJS2 (AJS scores at T2)[5]. It is hypothesised that LVs at the left hand side of the models affect LVs assigned to the right hand side in the path models. Conversely, LVs further to the left are not considered to be affected by LVs to the right in the path models.

The criterion variables are placed at the furthest right. Arrows directed from the LVs to MVs depict outward mode relationships and signify that MVs are reflectors of the corresponding LVs. In this study, the outward mode was employed, since the specified MVs were defined to reflect underlying dimensions of the corresponding LVs. Both the variables AJS1 and AJS2 were reflected by four MVs corresponding to the (a) *Composite factor (COMPOSIT1; COMPOSIT2); (b) Elements of art* factor (ELEMENTS1; ELEMENTS2); (c) *Formal properties* factor (FORMAL1; FORMAL2); and (c) *Symbolic level* factor (SYMBOL1; SYMBOL2).

The outer model results, or the information on the relationships between MVs and their corresponding LVs, are summarised in Table 2 for each grade level on the two assessment occasions, where only loadings of the MVs are given. It has been recommended that MVs should be excluded from the outer model if their loadings drop below ô0.40ô (Pedhazur 1982; Keeves 1992). Those MVs should be removed because they are believed to harm the 'predictive power' of the construct. However, with an intention to use the PLSPATH analyses results to identify possible patterns of relationships between variables on the separate occasions and across grade levels, it was appropriate that some MVs with loadings below this cut-off point be retained to keep the outer model equivalent for each grade level at two occasions. The results provided in Table 3 show the direct and total effects on the criterion variables AJS1 and AJS2, which served to evaluate the inner model.

Path models presented in Figures 1 to 8 advance the hypothesis that student performance in art production was affected both directly and indirectly by several other constructs of which GENDER, STUDAGE, ETHNIC and SES were exogenous variables or antecedents, since they were not influenced by other LVs. The remaining constructs were regarded as endogenous either because they were influenced by, or mediated the effects of other LVs. Attitude measures obtained on the VAAS and ASVASS in addition to teacher ratings of scholastic attitude were represented by the MVs VAHCM, VASAHCM and ATTITUT respectively, and were grouped together to form the LV ATTIT. The variable ASPIRAT (concerning students' expected occupations) originated from the Home Background Questionnaire and was employed to reflect the attitudinal dimension at the student level. It was assumed that ATTIT was an appropriate LV to be included with ASPIRAT in the related dimension. These variables were hypothesised to affect all subsequent LVs involving performance. Domain-relevant abilities were represented by the latent variables VISPACE, VAST and CREATE. The construct VISPACE was reflected by the five MVs: IDENTOTAL (*Identical Pictures Test* scores); CARDTOTAL (*Card Rotation Test* scores); HIDTOTAL (*Hidden Pictures Test* scores); RAWGEST (*Gestalt Completion Test* scores); and RAWVISUAL (*Surface Development Test* scores).

This LV in turn was hypothesised to influence the constructs VAST and CREATE. The relationship between intelligence (of which visual spatial ability is a component) and creativity has been well documented in previous research studies (Barron 1961; Carrol 1993; Getzels & Jackson 1962; Schubert 1973; Wallach 1971). The LV VAST was reflected by scores obtained on the *Visual Aesthetic Sensitivity Test* expressed through the MV RASCVAST. Other findings have emphasised the relative independence of VAST scores to student level

[5] T1 and T2 are used to describe results yielded from the analyses of AJS1 and AJS2 models respectively. The AJS1 path model employs AJS1 as the criterion, whereas the AJS2 path model employs AJS2 as the criterion. All MVs in these models involve measures obtained on the first occasion of test administration, with the exception of the criterion variable AJS2 that is formed from measures obtained on the second assessment occasion.

factors such as age, gender, ethnicity, intelligence and art training (Eysenck & Hawker 1994; Frois & Eysenck 1995; Iwawaki et al. 1979). The ordering of VAST within the causal hierarchy was intentional, for the purpose of testing the independence of its measures, as well as to investigate the predictive validity of the T factor (the good taste factor) on performance in art production. Similarly, the construct CREATE was reflected by the four MVs involving measures of figural creativity, namely ORIGA (originality), ELABORA (elaboration), FLUENCYA (fluency) and FLEXIBA (flexibility) obtained on the *Torrance Test of Creative Thinking.* These domain-relevant abilities were hypothesised to have an impact on student overall scholastic achievement, in addition to performance in art production. Scholastic achievement was represented by the latent variable ACHIEVE, which in turn was formed by a single MV, GPA (aggregate scores). Both direct and indirect effects of all predictors on the criterion variables, as well as effects involving interrelationships among other LVs in the model were examined.

Discussion

Outer Model

The loading pattern in each block of the outer models (Table 2) indicates the importance of contributing MVs to their respective LV. The variables GENDER, STUDAGE, STUDSHIP, OPPORT, VAST and ACHIEVE were estimated using unity mode, denoting each LV and its respective MV had a correlation of unity. Further examination of Table 2 reveals that a similar pattern of results was recorded for the outer models per grade level on each assessment occasion, with the exception of manifest variable EXPOCC2. This MV involved measures of students' expected occupation coded on a scale of socio-economic status (Broom et al. 1965). It was found that EXPOCC2 did not have a significant contribution for the Grade 9 sub-sample, and consequently it was deleted from the outer model.

Inner Model

Structural dimension

The inner model results presented considerable stability across grade levels and between occasions in the identification of direct or mediated effects between hypothesised student level factors. A slight variation was noted between grade levels and across occasions after examining the inner model effects (see Table 3) on the criteria AJS1 and AJS2.

Gender differences (GENDER)

The exogenous variable GENDER was shown to have an indirect and negative effect on AJS1 for all grade levels. The criterion AJS2 received positive direct effects from the exogenous variable GENDER for Grade 10, although GENDER influenced the criterion at T2 negatively both directly and indirectly for Grades 8 and 9, while also having an indirect negative effect for Grade 11.

With respect to the impact of student gender on the criterion variable, these findings reveal that females outperformed males in art production between occasions as well as across grade levels, except at Grade 10. Females' superior performance in art production in addition to their more favourable expressed attitudes, may indicate the effects of an underlying gender-role ideology and gender typing of school.

Furthermore, significant gender differences in spatial abilities were observed favouring the boys that agreed with previous research findings (Maccoby & Jacklin 1974; Fennema & Sherman 1977; Hyde 1981; Petersen 1976; Petersen et al. 1989; Rosenthal & Rubin 1982; Fennema & Tartre 1985). Grade 8 and Grade 9 data analyses point to boys' superior performance (Figures 1-4), although student gender failed to influence significantly

performance on the visual spatial abilities tests at Grades 10 and 11 (Figures 5-8). This result suggests that gender differences with respect to visual spatial abilities observed at earlier grade levels favouring boys, dissipated at Grade 10 possibly as a consequence of selection bias. These results are further supported by other studies, where age and gender differences on a test of spatial organisation were observed when administered to children and young adolescents, yet unconfirmed with adults (Kirk 1992). Moreover, meta-analyses on spatial visualisation ability showed gender differences to be in decline and close to extinction (Linn & Petersen 1985).

Ethnic background (ETHNIC)

The latent variable ETHNIC had a direct negative impact on performance in art production for Grade 8 on both assessment occasions. The inner model effects revealed a superior performance by students of non-English speaking background at this grade level. A positive and direct effect though, was identified from ETHNIC on the criterion variable AJS1 for Grade 10, in conjunction with a positive indirect effect being observed operating at the senior grade levels, implying students of Australian background outperformed non-English speaking background students.

Student age (STUDAGE)

In like manner, student age positively affected performance at Grade 9 with direct effects seen at both occasions, but an opposite effect held for the remaining models where negative indirect effects were recorded.

SHIP versus non-SHIP students (STUDSHIP)

Direct, positive effects of the construct STUDSHIP on the criterion variable at both T1 and T2, indicated that SHIP students outperformed non-SHIP students on AJS1 at Grades 8 and 11, as well as on AJS2 at Grades 10 and 11. However, additional positive indirect effects were identified from STUDSHIP on both AJS1 and AJS2 for the remaining models, verifying higher levels of performance in art production of SHIP students over their non-SHIP counterparts. This finding could imply a link between higher levels of intelligence and the acquisition of advanced techniques and skills for superior artistic performance (Luca & Allen 1974; Schubert 1973; Vernon, Adamson & Vernon 1977).

Attitudinal dimension of the home

Family involvement (FACTIVE)

The construct denoting family involvement in the visual arts (FACTIVE) had a direct, positive effect on Grade 11 students' performance in art production at T1, as well as positive indirect effects on the criterion at both assessment occasions operating through ATTIT. This implies that positive family attitudes contribute significantly to children's favourable deportment and overall performance in art production.

Process dimension of the home

Learning provision from the home (LPROVISI)

Learning provision in art from the home also had a small positive, indirect effect on AJS at the first assessment occasion for Grades 8 and 9, and somewhat larger indirect effects on performance at T2 for the senior grade levels mediated through student effort. With the exception of Grade 10, where learning provision was identified as a significant predictor of performance in art production at T1, no other direct effects were recorded on performance for the remaining grade levels. It is worth stating that learning provision was a construct formed by

observed variables measuring parental occupation coded on the Occupation Creative Scale (Helson et al. 1995). Other studies have shown that a cue-rich home environment may contribute to creativity-relevant skills (Ward 1969; Winner, Rosenstiel & Gardner 1976; Friedman, Raymond & Feldhusen 1978; Carother & Gardner 1979; Taunton 1984) and that students with art-related backgrounds, exhibit an increasing sensitivity to expressiveness promoted through art instruction (Bezruczko & Schroeder 1996). Since parents' expectations and reinforcement practices have been shown in the past to modify children's performance on intellectual-relevant behaviours (Williams 1976), it is not surprising that the learning provision in art at home was found to have some influence on performance in art production.

Attitudinal dimension at the student level

Students attitudes (ATTIT)

Overall, students' attitudes and aspirations positively affected performance both directly and indirectly, thereby stressing the importance of the attitudinal dimension at the student level in affecting the criterion variable. In addition, results (Figures 1-8) consistently showed that student attitudes had a small but positive relationship with performance on visual spatial abilities and visual aesthetic sensitivity measures. Scholastic achievement was also positively associated with student attitudes.

Student attitudes were shown to be influenced by six latent variables including gender, socio-economic status, SHIP membership, family activeness and future occupational aspirations. These variables were identified in the inner model obtained from the PLS path analyses. The findings suggested students expressing more favourable attitudes originated from home environments with an active interest and affinity toward the visual arts. Hence, parental attitudes were strong antecedents of children's attitudes. Furthermore, students of high socio-economic status homes expressed more positive affect, while students' attitudinal trends followed their occupational aspirations, with those aspiring for more creative occupations, professing more positive attitudes. PLS path analyses showed that girls perceived their parents as having more favourable attitudes toward their involvement in art education than boys, while parents from high SES homes were shown to be personally interested and to hold more favourable attitudes toward the study of the visual arts. Parental interests and feelings also influenced the extent of personal involvement and active contribution to children's art education. The findings supported the idea that parents of higher socio-economic status were more likely to be personally appreciative of the visual arts, providing greater opportunities for their children in that domain. The findings were consistent to a certain extent with those obtained in a study reporting an examination of the home environment and its relationship to attitude toward school, educational aspirations, career expectations and achievement of high school students in Cyprus (Koutsoulis 1995). In this study, Koutsoulis (1995) found that family processes influenced home environment more so than socio-economic status, particularly for females. More importantly, the home was found to affect students' academic self-concept and attitude toward school, while attitude also had effects both on achievement and future aspirations.

Process dimension at the student level

Student effort (EFFORT), opportunity to learn (OPPORT), and participation in art related activities (PARTICI)

EFFORT and OPPORT qualified as positive predictors of student performance for AJS1 and AJS2, revealing both direct and mediated positive effects on the criterion variables at all grade levels and across occasions. Therefore, the amount of time invested in art production and in receiving out-of-school instruction may have affected the outcome variables AJS1 and AJS2. Contrary to this, participation in extra-curricular activities presented small, positive, indirect

effects on AJS1 for Grade 8 and Grade 10, and on AJS2 for Grades 9, 10 and 11 operating through EFFORT.

Furthermore, results of the PLS path analyses indicated a consistent, positive and direct effect of student effort on the construct of figural creativity, across all grade levels. Since effort may be regarded as an indication of students' motivation towards the target behaviour, it is assumed that these findings point to a positive relationship between figural creativity and motivation to invest time and energy in visual art production and related homework. There is some overlap between this observation and Golann's (1962) interpretation of results obtained on a series of experiments regarding the self-actualising view of motivation for creativity. Golann (1962, p. 598) purported that "… creativity motivation is correlated with, but not sufficient for, the production of creative products", and that the motive would be reflected in preferences for stimuli or experiences which allow for self-expression. It is not surprising therefore, that an association between student effort (indication of motivation to be involved in art production and its related homework) and measures of figural creativity were identified in this study. There were no such effects observed though, from effort or opportunity to learn on the constructs of visual spatial ability or visual aesthetic sensitivity.

Domain-relevant abilities

Visual Spatial Abilities (VISPACE)

With respect to domain relevant abilities, VISPACE had a direct and positive influence on AJS in the analyses of Grades 8, 9 and 11 on T1 and Grades 8 and 10 data on T2. Also, an inspection of the total effects presented in Table 3 revealed that VISPACE positively influenced performance in art production on both occasions across all grade levels.

Visual Aesthetic Sensitivity (VAST)

VAST was shown to have direct and positive effects on AJS1 in the analyses of Grade 9 and Grade 10 data, as well as on AJS2 in the analyses of Grade 8, Grade 9 and Grade 11 data. Moreover, a positive and significant relationship was recorded between the constructs VAST and VISPACE.

Figural Creativity (CREATE)

A positive direct effect was observed from CREATE on AJS1 for Grades 8, 10, 11, and on AJS2 for Grades 9 and 11, but not Grade 8. In addition to the significant impact of visual spatial abilities on the VAST, examination of the inner model results (Table 3) revealed a direct and positive impact of visuospatial abilities on figural creativity. Congruent to this latter observation, Amabile (1996) in her work on the componential framework of creativity has argued that intelligence is a component of creative ability, identifying a positive relationship between measures of creativity and constructs of intelligence. Yamaha and Yu-Wen Tam (1996) have suggested that a person's intelligence has a predictive power for creative achievement, while higher intelligence is necessary for publicly recognised creative achievement. Moreover, some researchers (Wallach & Kogan, 1976) have contended that intelligence as well as creative thinking abilities, which are said not to be highly correlated with each other, might be necessary to attain creative achievements. This notion supports above average performance on measures of visuospatial abilities (constructs involving non-verbal intelligence) is almost a necessity for succeeding in creativity tasks (Clark & Zimmerman 1984a). The PLS analyses in this study further corroborate these reports.

Scholastic achievement (ACHIEVE)

Examination of the PLS path analyses results for T2 showed that the significant relationship between the two constructs (ACHIEVE and AJS2) was found particularly for Grade 10. Results of the analyses also indicated that scholastic achievement was related to performance on the visual spatial abilities tests, emphasising the positive relation between constructs of non-verbal intelligence and scholastic achievement.

CONCLUSION

PLSPATH results presented in this report have supported the construct validity of the AJS, by providing evidence of the interrelationship between the scale's measures with other behavioural measures, such as student level factors that assist in the meaningful interpretation of the observed response consistencies. The observation that different sets of factors influenced the criterion AJS at T1 as opposed to T2, could be attributed to a number of possible sources of variance, such as differences between the art works assessed at each occasion, the effects of instruction, or developmental changes occurring in particular dimensions. All in all, the criterion variable AJS was found to be directly affected by a number of latent variables such as figural creativity, visual spatial ability and visual aesthetic sensitivity. In addition, the findings demonstrated that constructs reflected by components of the AJS are not unrelated to the affective domain expressed in indicators such as student occupational intents and expressed attitudes, as well as identifying degree of effort that involves an expression of student motivation. It could be therefore supported that the Aesthetic Judgement Scale provides meaningful measures of students' performance in art production that carry practical implications for art education.

REFERENCES

Amabile, T.M. (1996). Creativity in context. Update to the Social Psychology of Creativity. Colorado, USA: Western Press.

Armstrong, C.L. (1994). Designing assessment in art. National Art Education Association. Reston VA.

Barron, F. (1961). Creative vision and expression in writing and painting. In Conference on the creative person, Berkeley: University of California, Institute of personality Assessment and Research.

Barron, F. (1972). Artists in the Making. New York: Seminar Press.

Barron, F. (1974). Basic Research in Aesthetic Education. Washington, D.C.: U.S. Office of Education.

Barron, F. (1977). Symbolic Scope as a Predictor of Ability in Art. In S.L. Carmean and B.L. Grover (Eds.), Creative Thinking, Eighth Western Symposium on Learning. Bellingham: Western Washington University, 1977.

Barron, F. (1982). Images - The universal language. New directions for testing and measurement. San Francisco, 3, 111-123.

Beattie, D.K. (1997). Visual arts criteria, objectives, and standards: A revisit. National Art Education Association. 38(4), 217-231.

Beittel, K.R., & Burkhart, R.C. (1963). Statements of spontaneous, divergent, and academic art students. Studies in Art Education, 5(1), 20-41.

Bezruczko, N., & Schroeder, D.H. (1996). The Development of visual preferences in art-trained and non-art-trained schoolchildren. Genetic, Social, and General Psychology Monographs, 122(2), 179-196.

Boston, N.E. (1987). Determining giftedness in elementary visual art students. South Bend, IN: Indiana University South Bend. (ERIC Document Reproduction Service No 301 025).

Broom, L., Jones, L.F., & Zubrzycki, J. (1965). An Occupational Classification of the Australian Workforce. Australian and New Zealand Journal of Sociology, 1, Supplement.

Burkhart, R.C. (1962). Spontaneous and deliberate ways of learning. Scanton, PA: International Textbook Co.

Carother, T., & Gardner, H. (1979). When children's drawings become art: The emergence of aesthetic production and perception. Developmental Psychology, 15, 570-580.

Carroll, J.B. (1993). Human cognitive abilities: A survey of factor analytic studies. New York, Cambridge University Press.

Chetelat, F.J. (1981). Visual arts education for the gifted elementary level art student. Gifted Child Quarterly, 25, 154-158.

Chin, W.W. & Newsted, P.R. (1999). Structural equation modelling analysis with small samples using partial least squares. In R.H Hoyle, Statistical strategies for small sample research. Ca: Sage Publications, 307-341.

Clark, G.A., & Zimmerman, E. (1984). Educating artistically talented students. Syracuse, NY: Syracuse University Press.

Clark, G.A., & Zimmerman, E. (1988). Views of self, family background, and school: Interviews with artistically talented students. Gifted Child Quarterly, 32(4), 340-346.

Clark, G.A., & Zimmerman, E. (1992). Issues and Practices Related to Identification of Gifted and Talented Students in the Visual Arts. Identification in the Arts. Research-Based Decision Making Series. No. 9202.

Cohen, J.A. (1941). A scale for the measurement of attitude toward the aesthetic value. Journal of Psychology, 12, 75-79.

Csikszentmihalyi, M., & Getzels, J.W., (1971). Discovery oriented behaviour and the originality of creative products. Journal of Personality and Social Psychology, 19, 47-52.

Ellam, AH, Godown, M., & Doughty, R. (1988). Guidelines for the identification of artistically gifted and talented students (revised). Columbia, SC: South Carolina State Department of Education. (ERIC Document Reproduction Service No ED 306 761).

Eysenck, H,J. & Hawker, G.W. (1994). The taxonomy of visual aesthetic preferences: an empirical study. Empirical Studies of the Arts, 12(1), 95-101.

Fennema, E.,& Sherman, J. (1977). Sex related differences in mathematics achievement, spatial visualisation, and affective factors. American Educational Research Journal, 14, 51-71.

Fennema, E., & Tartre, L.A. (1985). The use of spatial visualisation in mathematics by boys and girls. Journal of Research in Mathematics and Education, 16, 184-206.

French, J.W., Ekstrom, R.B., & Price, L.A. (1963). Manual and kit of reference tests for cognitive factors. Princeton, NJ: Educational Testing Service.

Friedman, F., Raymond, B.A., & Feldhsen, J.F. (1978). The effects of environmental scanning on creativity. Gifted Child Quarterly, 22, 248-251.

Frois, J.P., & Eysenck, H.J. (1995). The Visual Aesthetic Sensitivity Test applied to Portuguese children and fine arts students. Creativity Research Journal. 8(3), 277-284.

Getzels, J.W. , & Jackson, P.W. (1962). Creativity and intelligence: Explorations with gifted students. New York: Wiley.

Ghallagher, S.A. (1985). A comparison of the concept of over-excitability with measures of creativity and school achievement in sixth grade students. Roeper Review, 8(2), 115-119.

Götz, K.O. (1985). VAST: Visual Aesthetic Sensitivity Test. (4th Ed.) Dusseldorf, Germany: Concept Verlag.

Hardiman, G.W., Liu, F.J., & Zernich, T. (1992). Assessing knowledge in the visual arts. In G.C. Cupchik, & J. Laszlo (Eds.), Emerging visions of the aesthetic process. Psychology, semiology, and philosophy. Cambridge University Press, USA.

Helson, R., Roberts, & Agronick, G. (1995). Enduringness and change in creative personality and the prediction of occupational creativity. Journal of Personality and Social Psychology, 69(6), 1173-1183.

Hurwitz, A., & Day, M. (1991). Children and their art: Methods for the elementary school (5th Edition). San Diego, CA: Harcourt, Brace, Jovanovich.

Hyde, J.S. (1981). How large are cognitive gender differences? A meta-analysis. Psychological Bulletin, 104, 53-69.

Iwawaki, S., Eysenck, H.J., & Götz, K.O. (1979). A new Visual Aesthetic Sensitivity test (VAST): II. Cross cultural comparison between England and Japan. Perceptual and Motor Skills, 49, 859-862.

Jellen, H.G., & Verduin, J.R. (1986). Handbook for differential education of the gifted: A taxonomy of 32 key concepts. Carbondale, IL: Southern Illinois University Press.

Kaplan, A (1964). The Conduct of Inquiry. San Francisco, CA: Chandler

Keeves, J.P. (1992). The IEA study of Science III. Changes in science education and achievement: 1970 to 1984. Oxford: Pergamon Press.

Kirk, U. (1992). Evidence for early acquisition of visual organisation ability: A developmental study. The Clinical Neuropsychologist, 6(2), 171-177.

Koutsoulis, M.K. (1995). Home environment and its relationship to self-concetp, attitude toward school, educational aspirations, career expectations, and achievement of high school students in Cyprus. Dissertation Abstracts International Section A: Humanities and Social Sciences, 56(6-A): 2194.

Lietz, P. (1996). Changes in reading comprehension across culture and over time. M_nster: Waxmann.

Linn, M.C., & Petersen, A.C. (1985). Emergence and characterisation of sex differences in spatial abiltiy: A meta-analysis. Child development, 56, 1479-1498.

Luca, M., & Allen, B. (1974). Teaching gifted children art in grades one through three. Sacramento, CA: California State Department of Education. (ERIC Document Reproduction Service No. ED 082 433).

Maccoby, E.E. & Jacklin, C.N. (1974). The Psychology of sex differences, Stanford, CA: University Press.

Pedhazur, E.J. (1982). Multiple regression in behavioural research (2nd ed.). New York: Holt, Rinehart and Winston.

Petersen, A.C. (1976). Physical androgyny and cognitive functioning in adolescence. Developmental Psychology, 12, 524-533.

Petersen, N.S., Kolen, MJ., Hoover, H.D. (1989). Scaling, Norming, and Equating. In: Linn, R.L. (Ed), Educational Measurement (3rd edn). American Council on Education and Macimillan Publishing Company, New York, pp. 221-262.

Renzulli, J.S. (1982). The enrichment triad. In C.J. Maker, (Ed.), Teaching models I education of the gifted (pp. 207-236). Austin, TX: Pro-ed.

Renzulli, J.S., Reis, S.M., & Smith, L.H. (1981). The revolving door identification model (RDIM). Mansfield Centre: Creative Learning Press.

Renzulli, J.S., & Smith, L.H. (1977). Two approaches to the identification of gifted students. Exceptional Children, 43, 512-518.

Rosenthal, R., & Rubin, D.B. (1982). Further meta-analytic procedures for assessing cognitive gender differnces. Journal of Educational Psychology, 74, 708-712.

Saunders, R.J. (1982). Screening and identifying the talented in art. Roeper Review, 4(3), 7-10.

Schubert, D.S.P. (1973). Intelligence as necessary but not sufficient for creativity. Journal of Genetic Psychology, 122, 45-47.

Scott, L.E. (1988). A comparative study of personality, values, and back-ground characteristics of artistically talented, academically talented, and average 11th and 12th grade students. Studies in Art Education, 29(3), 292-301.

Sellin, N. (1990). PLSPATH Version 3.01. Program Manual. Hamburg, West Germany.

Silance, E.B., & Remmers, H.H. (1934). An experimental generalised master scale: A scale to measure attitudes toward any school subject. Purdue Univ. Stud. high. Educ., XXVI, 35, 84-88.

Taunton, M. (1984). Four-year-old children's recognition of expressive qualities in painting reproductions. Journal of Psychology, 57, 239-242.

Torrance, E.P. (1974). Torrance tests of creative thinking: Norms-technical manual (research edition). Princeton, NJ: Personnel Press.(research edition). Princeton, NJ: Personnel Press.

Vernon, P.D., Adamson, G., & Vernon, D. (1972). The psychology and education of gifted children. Boulder, CO: Viewpoint Press.

Wallach, M.A. (1971). The creativity-intelligence distinction. New York: General Learning Press.

Wallach, M.A., & Kogan, N. (1976). Creativity and intelligence in children. In A. Rothenberg & C.R. Hausman (Eds.), The creativity question (pp.208-217). Durham, NC: Duke University Press. (Original work published 1972).

Ward, W.C. (1969). Creativity and environmental cues in nursery school children. Developmental Psychology, 1, 543-547.

Wenner, G.C. (1985). Discovery and recognition of the artistically talented. Journal for the Education of the Gifted, 8(3), 221-238.

Williams, T. (1976). Abilities and environments. In W.H. Sewell, R. M. Hauser and D.L. Featherman (eds.) Schooling and Achievement in American Society, NY: Academic Press.

Winner, E., Rosenstiel, A.K., & Gardner, H. (1976). Development of metaphoric understanding. Developmental Psychology, 12, 289-297.

Wold, H. (1982). Soft modeling. The basic design and some extensions. In K.G. Jöreskog & H. Wold (Eds.). Systems under indirect observation. Part III. Amsterdam, The Netherlands: North Holland Press.

Yamaha, H., & Yu-Wen Tam, A. (1996). Prediction Study of Adult Creative Achievement: Torrance Longitudinal Study of Creativity Revisited. (Research Note). Journal of Creative Behaviour, 30(2), 144-149.

Table 1. Summary of latent variables (LV) and manifest variables (MV)

Dimensions	Latent variables	Manifest variables
Criterion	17. AJS1 – occasion 1	Scores on samples of work with the AJS: COMPOSIT1- Score on Composite factor ELEMENTS1- Score on Elements of art factor FORMAL1- Score on Formal properties factor SYMBOL1- Score on Symbolic level factor
	17. AJS2 – occasion 2	Scores on samples of work with the AJS: COMPOSIT2 - Score on Composite factor ELEMENTS2 - Score on Elements of art factor FORMAL2 - Score on Formal properties factor SYMBOL2 - Score on Symbolic level factor
Scholastic Achievement	16. ACHIEVE	Student grade point average (GPA)
Domain -relevant abilities	15. CREATE	Scores on Torrance Creative thinking test ORIGINA - Score on Originality ELABORA - Score on Elaboration FLUENCYA - Score on Fluency FLEXIBA - Score on Flexibility
	14. VAST	RASCVAST - Scores on the Visual Aesthetic Sensitivity test
	13. VISPACE	CARDTOTAL - Scores on Card rotation test RAWGEST - Scores on Gestalt completion test RAWVISUAL - Scores on Surface development test IDENTOTAL - Scores on Identical pictures test HIDTOTAL - Scores on Hidden figures test
Process at student level	12. EFFORT	SPEND - Student spends time after school on art production TIME - Amount of time spent on art production MARKS - Marks received on art production
	11. OPPORT	Opportunity to learn PRVS - Student has attended art classes after hours in the past
	10. PARTICI	Student participation in art related activities APROG - Watches art programs on TV READ - Reads books on the visual arts EXHIB - Visits art exhibitions and galleries LIBRA - Visits the library to study issues related to the visual arts
Attitudinal at student level	9. ATTIT	VAAHCM - Score on Attitude toward Visual Arts scale VASAHCM - Score on Attitude toward Visual Arts Subject Scale ATTITUT - Score on General Scholastic Attitude toward all subjects
	8. ASPIRAT	Student aspiration EXPOCC1- Expected occupation measured on a scale for degree of creativity EXPOCC2 - Expected occupation measured on a scale for socio-economic status
Process of the home	7. LPROVISI	Learning provision PINVOL - Parental involvement in the visual arts RINVOL - Relative involvement in the visual arts VISIT - Visits to countries of rich cultural heritage PHELP - Parental help on homework relating to the visual arts subject MOCC1- Mother's occupation measured on a scale of degree of creativity FOCC1- Father's occupation measured on a scale of degree of creativity
Attitudinal of the home	6. FACTIVE	Family activeness PFEEL - Parent's feeling regarding student studying the visual art PNTER - Parental interest in the visual arts
Structural	5. STUDSHIP	SHIP - SHIP or non-SHIP category
	4. SES	Socio-economic status including: RESOU - Resources found at home involving the visual arts such as books, magazines, videos, aesthetic objects etc. FOCC2 - Father's occupation measured on the occupational scale MOCC2 - Mother's occupation measured on the occupational scale FEDUC - Father's education MEDUC - Mother's education
	3. ETHNIC	CBIRTH - Student's country of birth FBIRTH - Father's country of birth MBIRTH - Mother's country of birth LANGU - Language spoken at home LIVAU - Years residing in Australia
	2. STUDAGE	AGE - Student age in months
	1. GENDER	STUDSEX - Student gender

Table 2. Summary of Loadings for Manifest Variables.

	Occasion 1				Occasion 2			
Variable	**Grade 8**	**Grade 9**	**Grade 10**	**Grade 11**	**Grade 8**	**Grade 9**	**Grade 10**	**Grade 11**
GENDER								
STUDSEX	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
STUDAGE								
AGE	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ETHNIC								
CBIRTH	0.87	0.93	0.92	0.83	0.90	0.93	0.81	0.82
FBIRTH	0.61	0.74	0.57	0.77	0.59	0.74	0.68	0.78
MBIRTH	0.81	0.80	0.43	0.67	0.79	0.81	0.63	0.68
LANGU	0.50	0.90	0.49	0.88	0.65	0.82	0.61	0.88
LIVAU	0.85	0.83	0.92	0.89	0.87	0.91	0.86	0.87
SES								
RESOU	0.80	0.72	0.56	0.58	0.80	0.73	0.56	0.71
FOCC2	0.35	0.53	0.58	0.78	0.35	0.53	0.58	0.74
MOCC2	0.54	0.65	0.64	0.65	0.54	0.64	0.63	0.57
FEDUC	0.55	0.52	0.63	0.71	0.55	0.52	0.63	0.68
MEDUC	0.58	0.62	0.72	0.58	0.58	0.62	0.72	0.51
STUDSHIP								
SHIP	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FACTIVE								
PNTER	0.74	0.77	0.76	0.84	0.74	0.77	0.75	0.82
PFEEL	0.89	0.84	0.83	0.87	0.89	0.84	0.85	0.88
LPROVISI								
PINVOL	0.70	0.49	0.69	0.78	0.70	0.50	0.69	0.80
RINVOL	0.56	0.52	0.57	0.73	0.56	0.52	0.58	0.75
PHELP	0.41	0.47	0.44	0.48	0.41	0.47	0.44	0.47
FOCC1	0.50	0.38	0.67	0.40	0.49	0.37	0.66	0.39
MOCC1	0.59	0.67	0.60	0.57	0.58	0.65	0.59	0.53
VISIT	0.52	0.52	0.48	0.28	0.52	0.53	0.49	0.30
ASPIRAT								
EXPOCC1	0.91	1.00	0.95	0.96	0.93	1.00	0.95	0.96
EXPOCC2	0.65	Deleted	0.55	0.48	0.62	Deleted	0.56	0.47

Table 11.2. Summary of Loadings for Manifest Variables (continued)

ATTIT								
VASAHCM	0.77	0.68	0.84	0.52	0.77	0.68	0.81	0.52
VAAHCM	0.71	0.70	0.75	0.82	0.72	0.70	0.72	0.83
ATTITUT	0.76	0.71	0.67	0.83	0.76	0.71	0.72	0.81
PARTICI								
APROG	0.78	0.75	0.78	0.62	0.77	0.75	0.78	0.62
READ	0.77	0.83	0.78	0.72	0.75	0.83	0.78	0.72
EXHIB	0.75	0.76	0.78	0.75	0.75	0.76	0.77	0.75
LIBRA	0.73	0.68	0.67	0.80	0.73	0.68	0.67	0.80
OPPORT								
PRVS	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
EFFORT								
SPEND	0.81	0.78	0.86	0.77	0.80	0.80	0.87	0.76
TIME	0.87	0.89	0.86	0.83	0.88	0.90	0.84	0.83
MARKS	0.56	0.68	0.61	0.66	0.57	0.66	0.63	0.67
VISPACE								
CARDTOTAL	0.78	0.74	0.80	0.73	0.78	0.72	0.81	0.74
RAWGEST	0.40	0.75	0.57	0.63	0.41	0.76	0.60	0.64
RAWVISUA	0.75	0.72	0.79	0.75	0.75	0.71	0.78	0.71
IDENTOTAL	0.66	0.73	0.60	0.76	0.66	0.73	0.60	0.78
HIDTOTAL	0.60	0.37	0.70	0.64	0.59	0.40	0.68	0.64
VAST								
RASCVAST	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CREATE								
ORIGA	0.78	0.91	0.89	0.90	0.83	0.91	0.90	0.91
ELABORA	0.92	0.79	0.68	0.79	0.87	0.79	0.49	0.74
FLUENCYA	0.55	0.81	0.49	0.89	0.62	0.77	0.63	0.90
FLEXIBA	0.71	0.90	0.90	0.92	0.78	0.88	0.95	0.94
ACHIEVE								
GPA	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AJS								
COMPOSIT	0.98	0.97	0.96	0.98	0.98	0.98	0.97	0.97
ELEMENTS	0.97	0.96	0.96	0.97	0.98	0.98	0.98	0.97
FORMAL	0.96	0.97	0.95	0.97	0.97	0.97	0.96	0.97
SYMBOL	0.96	0.96	0.95	0.96	0.98	0.97	0.98	0.96
N	201	87	96	67	201	87	96	67

Table 3. Inner model effects on AJS by grade level and occasion.

	Occasion 1					Occasion 2				
Direct Effects	**Grade 8**	**Grade 9**	**Grade 10**	**Grade 11**	**Total**	**Grade 8**	**Grade 9**	**Grade 10**	**Grade 11**	**Total**
GENDER	-	-	-	-	None	-0.13	-0.10	0.19	-	3
STUDAGE	-	0.30	-	-	1	-	0.20	-	-	1
ETHNIC	-0.10	-	0.23	-	2	-0.17	-	-	-	1
SES	-	-	-	-	None	-	-	-	-	-
STUDSHIP	0.08	-	-	0.16	2	-	-	0.12	0.20	2
FACTIVE	-	-	-	0.18	1	-	-	-	-	-
LPROVISI	-	-	0.12	-	1	-	-	-	-	-
ASPIRAT	0.13	0.16	-	-	2	0.08	0.27	-	-	2
ATTIT	0.31	-	0.18	0.17	3	0.33	-	-	0.40	2
PARTICI	-	-	-	-	None	-	-	-	-	-
OPPORT	-	0.15	-	-	1	-	0.17	-	-	1
EFFORT	0.12	-	-	-	1	-	-	0.36	0.15	2
VISPACE	0.09	0.27	-	0.13	3	0.10	-	0.18	-	2
VAST	-	0.14	0.09	-	2	0.08	0.21	-	0.23	3
CREATE	0.13	-	0.15	0.20	3	-	0.14	-0.24	0.11	3
ACHIEVE	-	-	-	-	None	-	-	0.32	-	1
Total Effects										
GENDER	-0.16	-0.04	-0.10	-0.10	4	-0.26	-0.22	0.05	-0.11	4
STUDAGE	-0.03	0.24	-0.02	-0.06	4	-0.02	0.14	-0.08	-0.09	4
ETHNIC	-0.10	0.05	0.24	0.06	4	-0.17	0.01	0.08	0.02	4
SES	0.10	0.07	0.17	0.15	4	0.07	0.09	0.22	0.18	4
STUDSHIP	0.22	0.03	0.03	0.16	4	0.15	0.03	0.17	0.26	4
FACTIVE	0.18	0.08	0.08	0.28	4	0.13	0.14	0.10	0.29	4
LPROVISI	-	-	0.13	-	1	-	-	0.02	0.02	2
ASPIRAT	0.17	0.19	0.04	0.03	4	0.12	0.30	0.01	0.12	4
ATTIT	0.37	0.18	0.20	0.19	4	0.37	0.13	0.29	0.42	4
PARTICI	0.01	-	0.01	-	2	-	0.01	0.07	0.05	3
OPPORT	0.02	0.18	0.01	-	3	-	0.17	0.01	-	2
EFFORT	0.12	-	0.03	-	2	-	0.03	0.36	0.15	3
VISPACE	0.09	0.27	0.04	0.21	4	0.11	0.07	0.26	0.08	4
VAST	-	0.14	0.09	-	2	0.08	0.21	-	0.23	3
CREATE	0.13	-	0.15	0.20	3	-	0.14	-0.24	0.11	2
ACHIEVE	-	-	-	-	None	-	-	0.32	-	1
% Variance Explained	29	17	20	25		27	22	40	38	

N	201	87	96	67		201	87	96	67

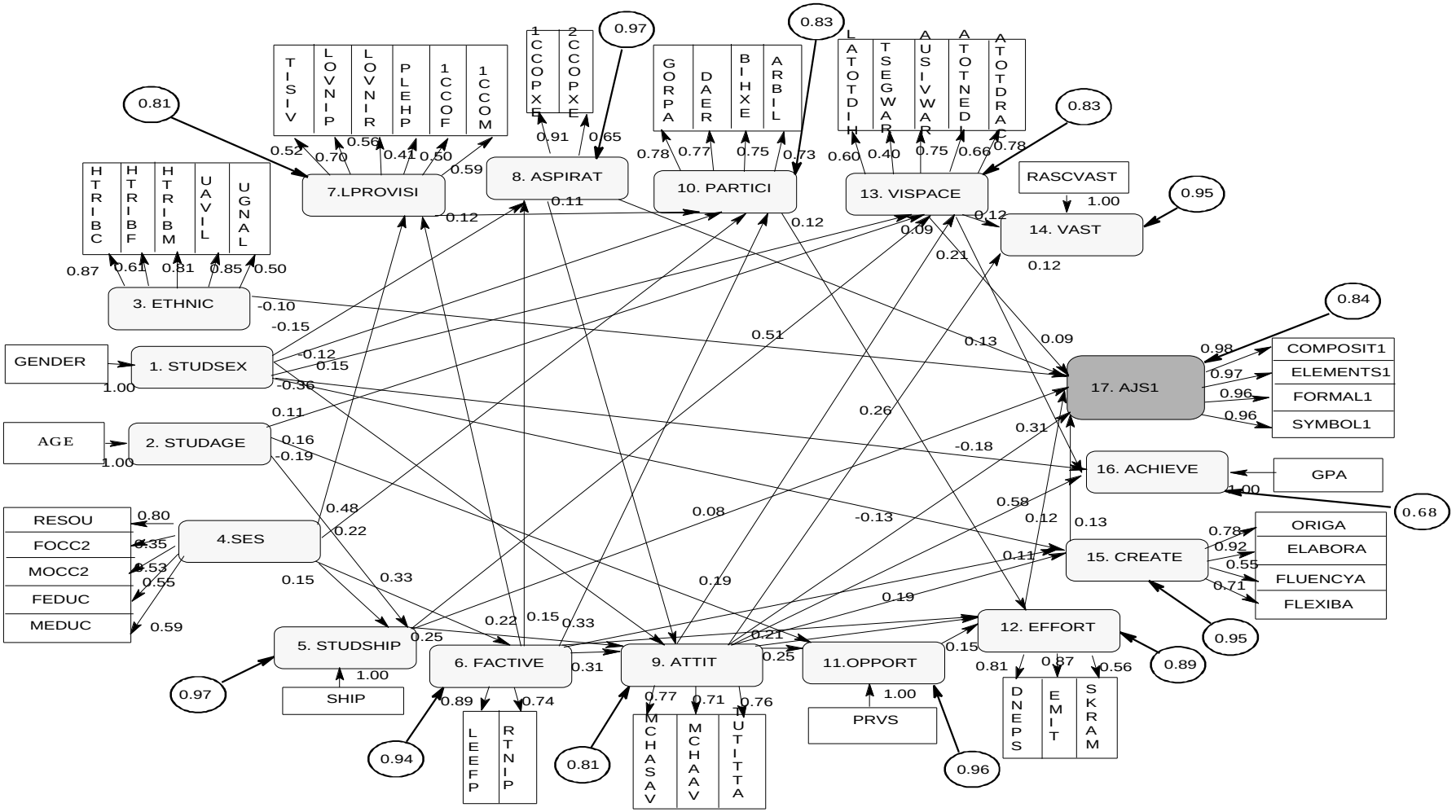


Figure 1. Grade 8 - Student level factors influencing performance in art production with AJS1 as the criterion variable.

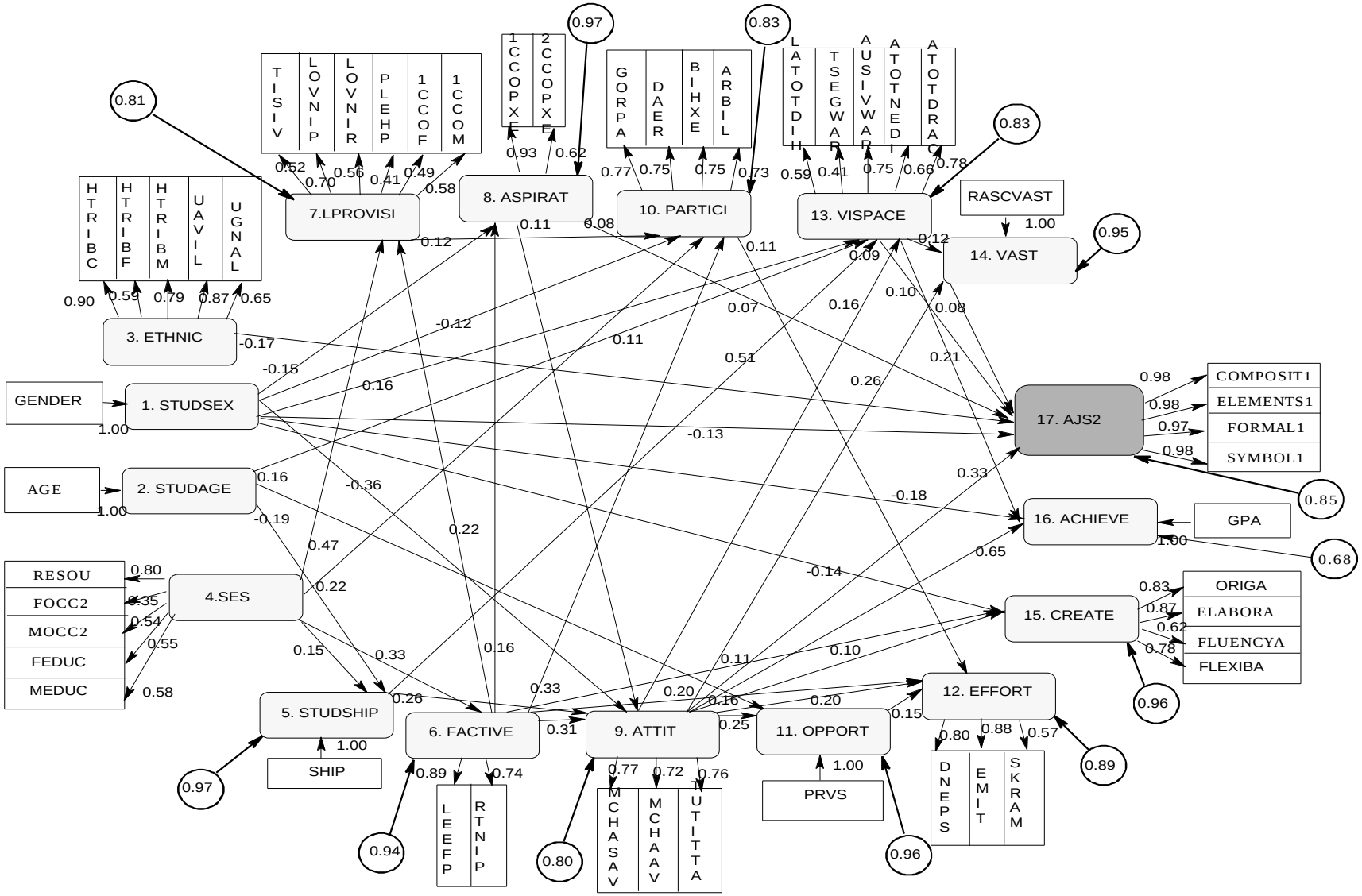


Figure 2. Grade 8 - Student level factors influencing performance in art production with AJS2 as the criterion variable.

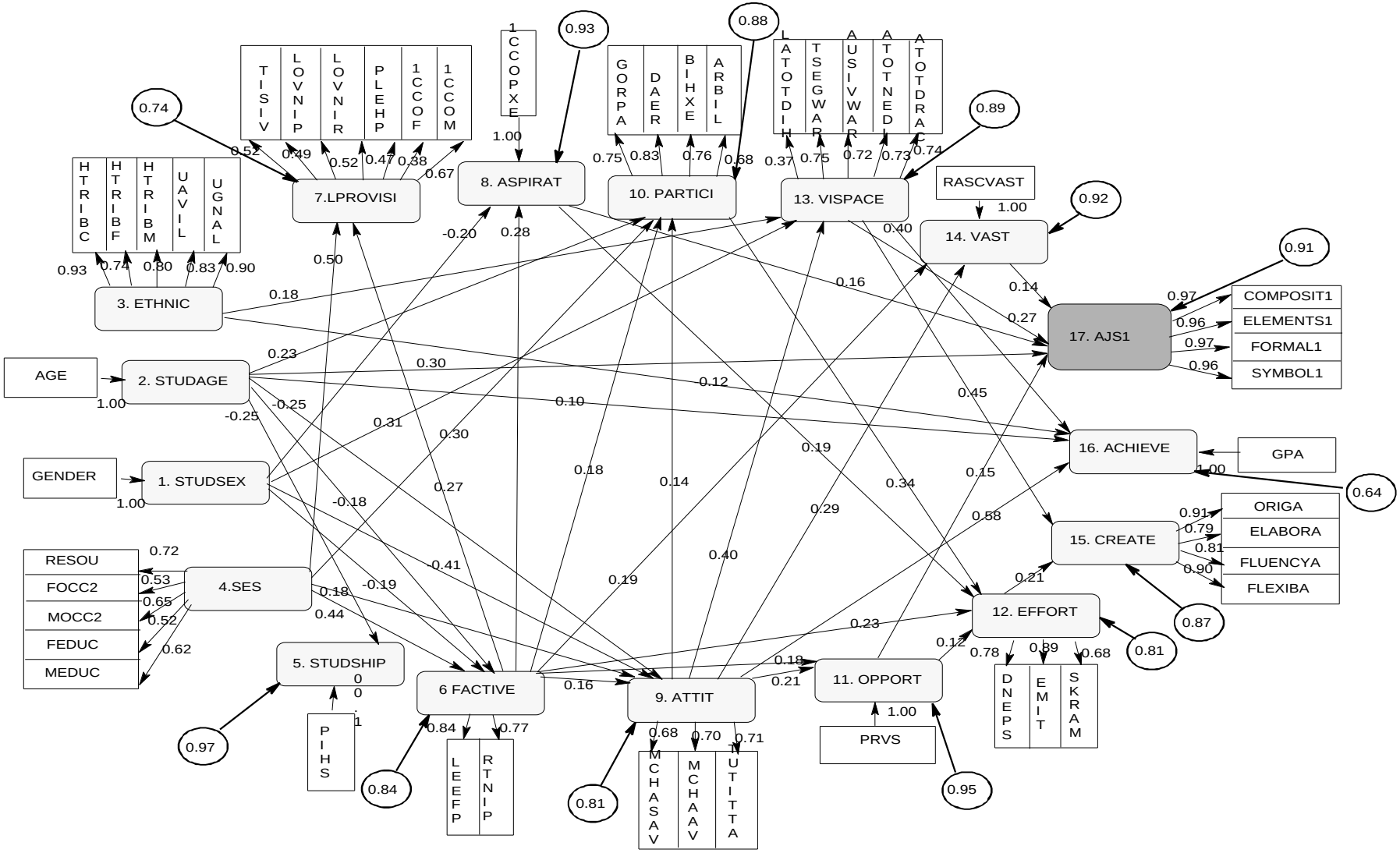


Figure 3. Grade 9 - Student level factors influencing performance on art production with AJS1 as the criterion variable.

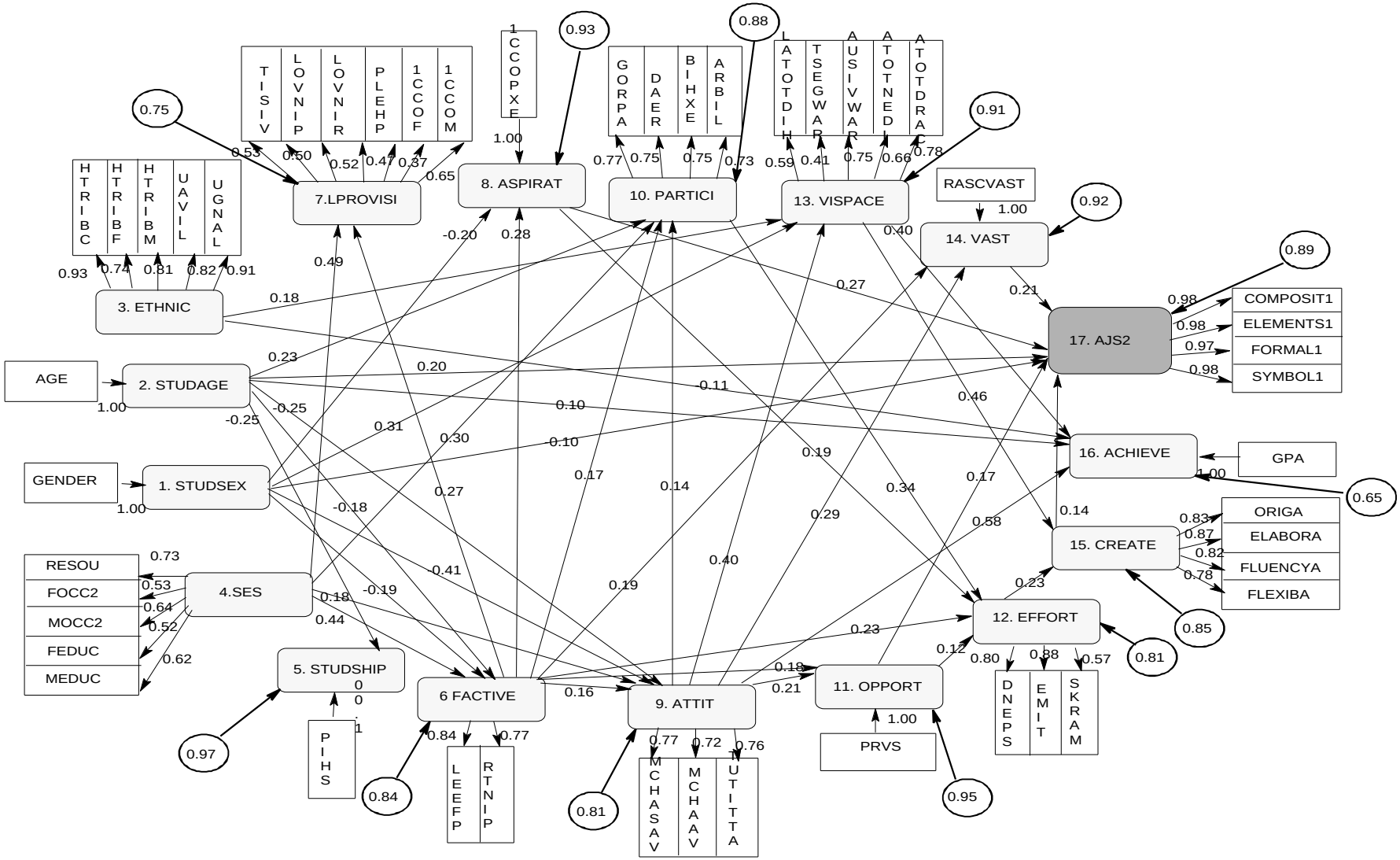


Figure 4. Grade 9 - Student level factors influencing performance on art production with AJS2 as the criterion variable.

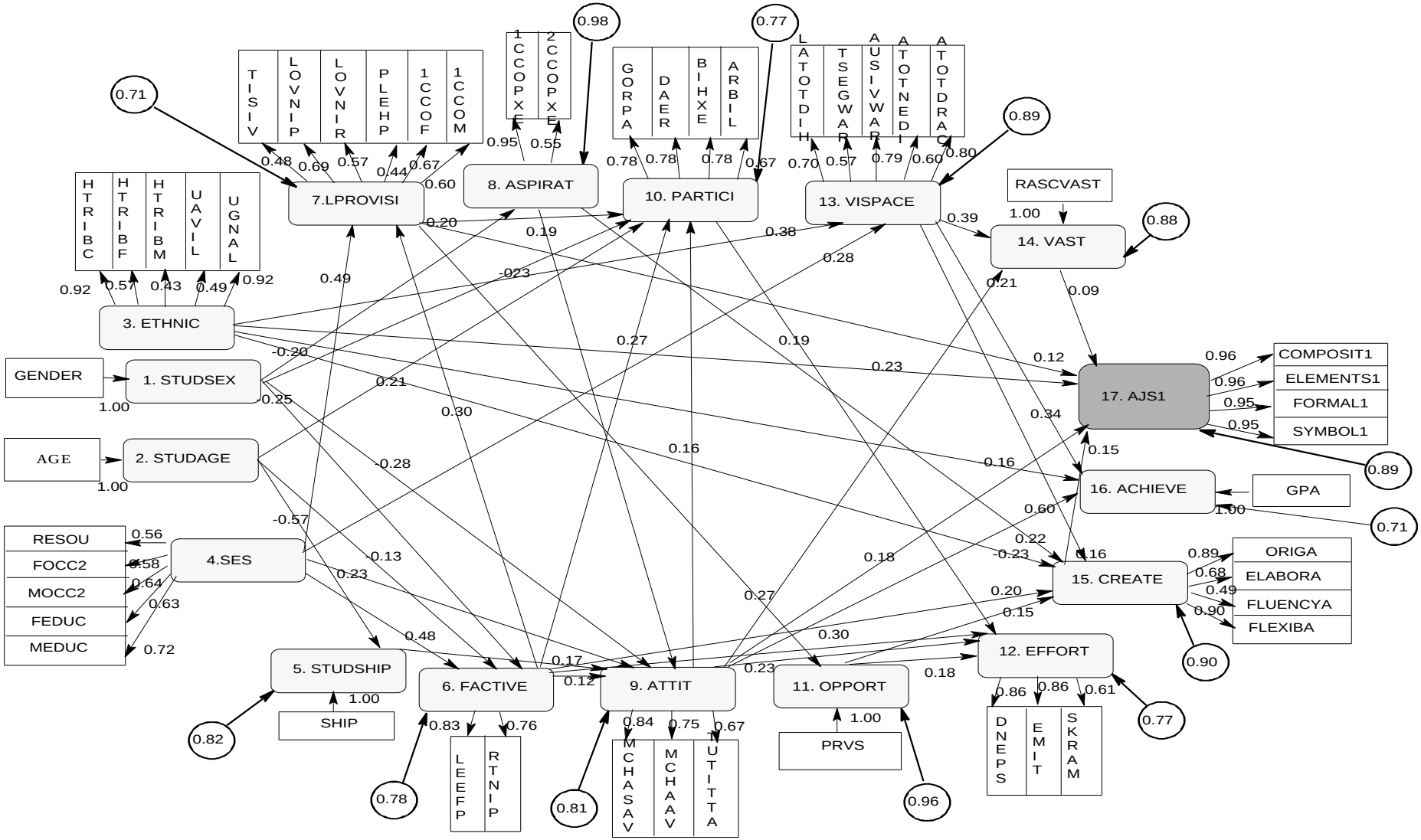


Figure 5. Grade 10 - Student level factors influencing performance on art production with AJS1 as the criterion variable.

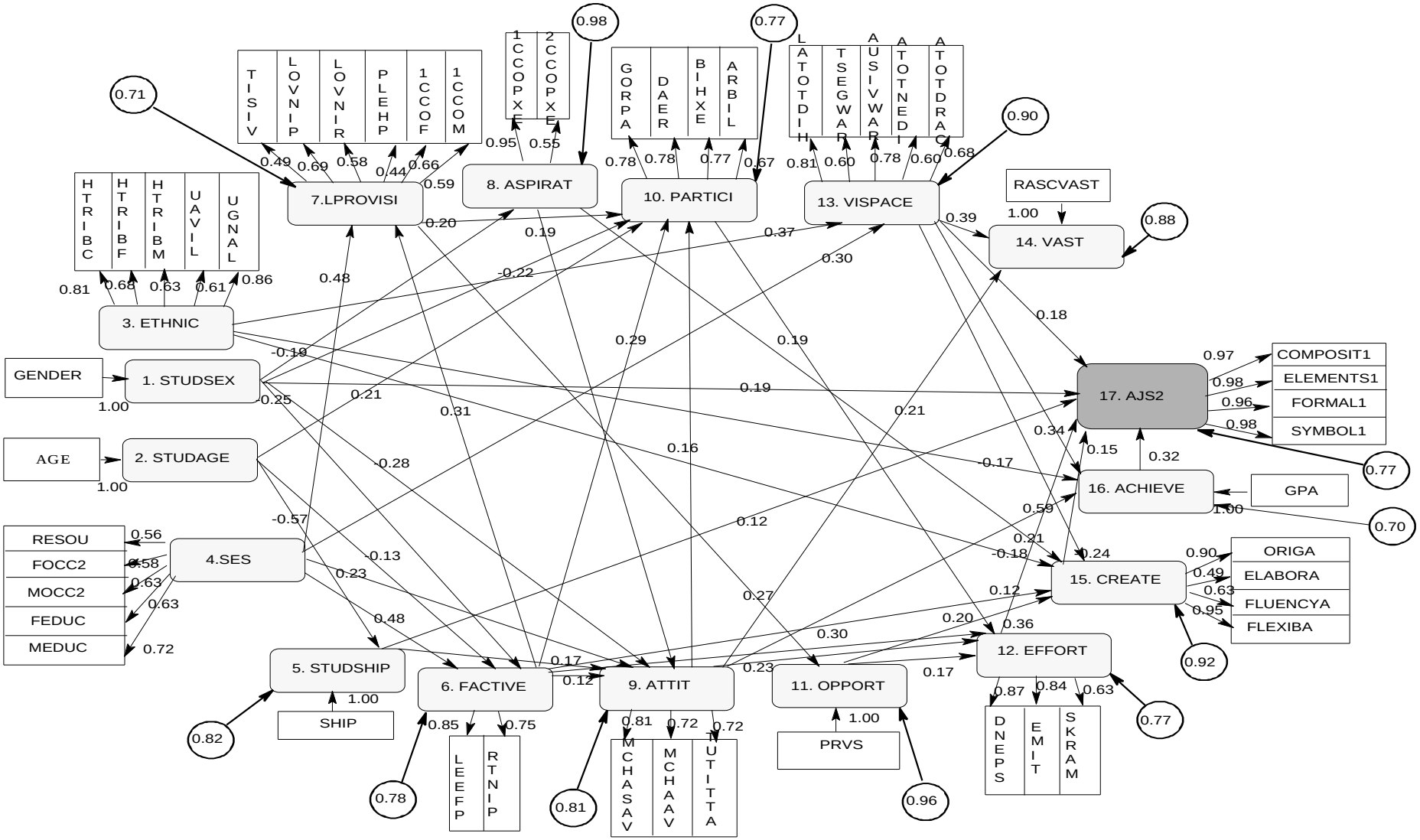


Figure 6. Grade 10 - Student level factors influencing performance on art production with AJS2 as the criterion variable.

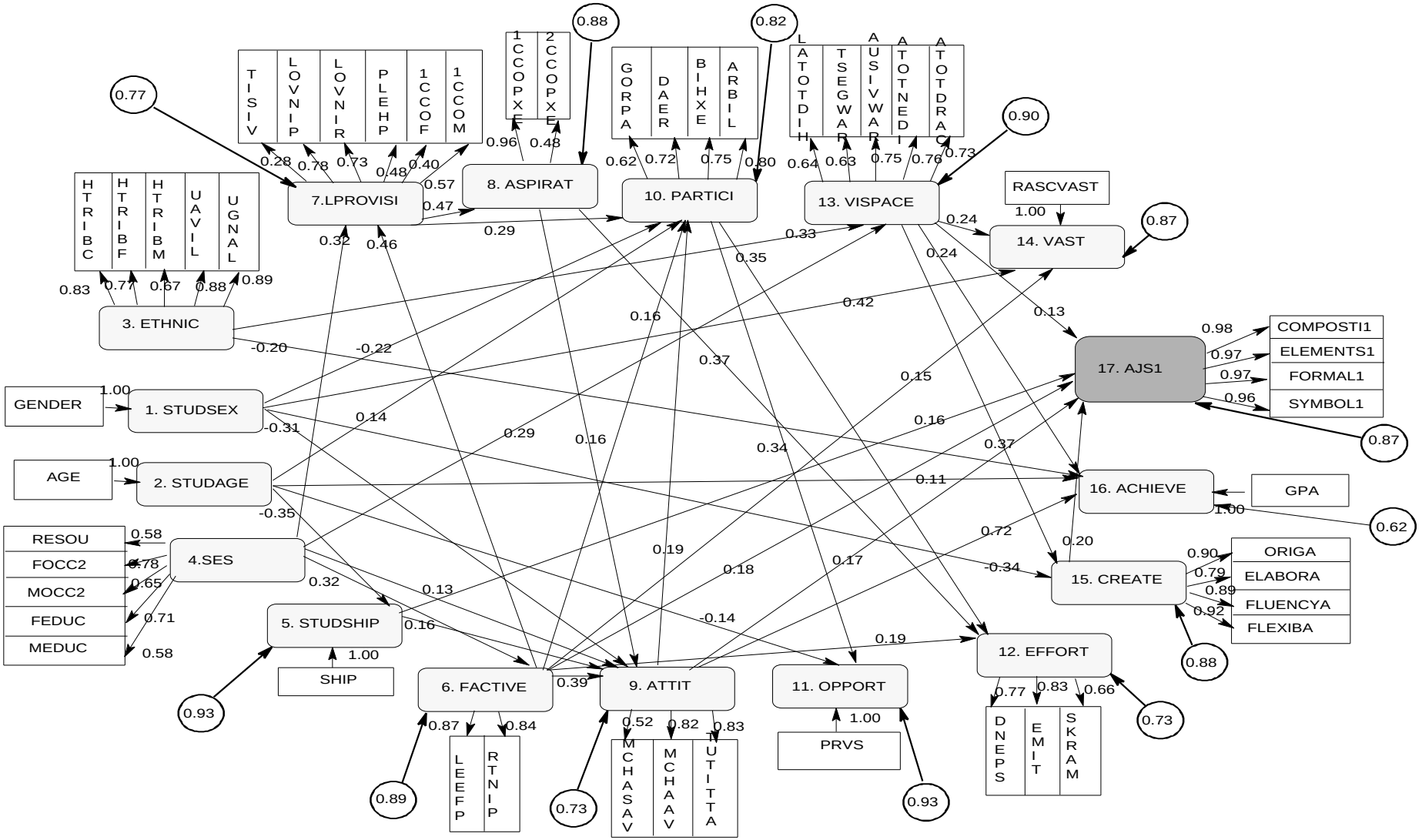


Figure 7. Grade 11 - Student level factors influencing performance on art production with AJS1 as the criterion variable.

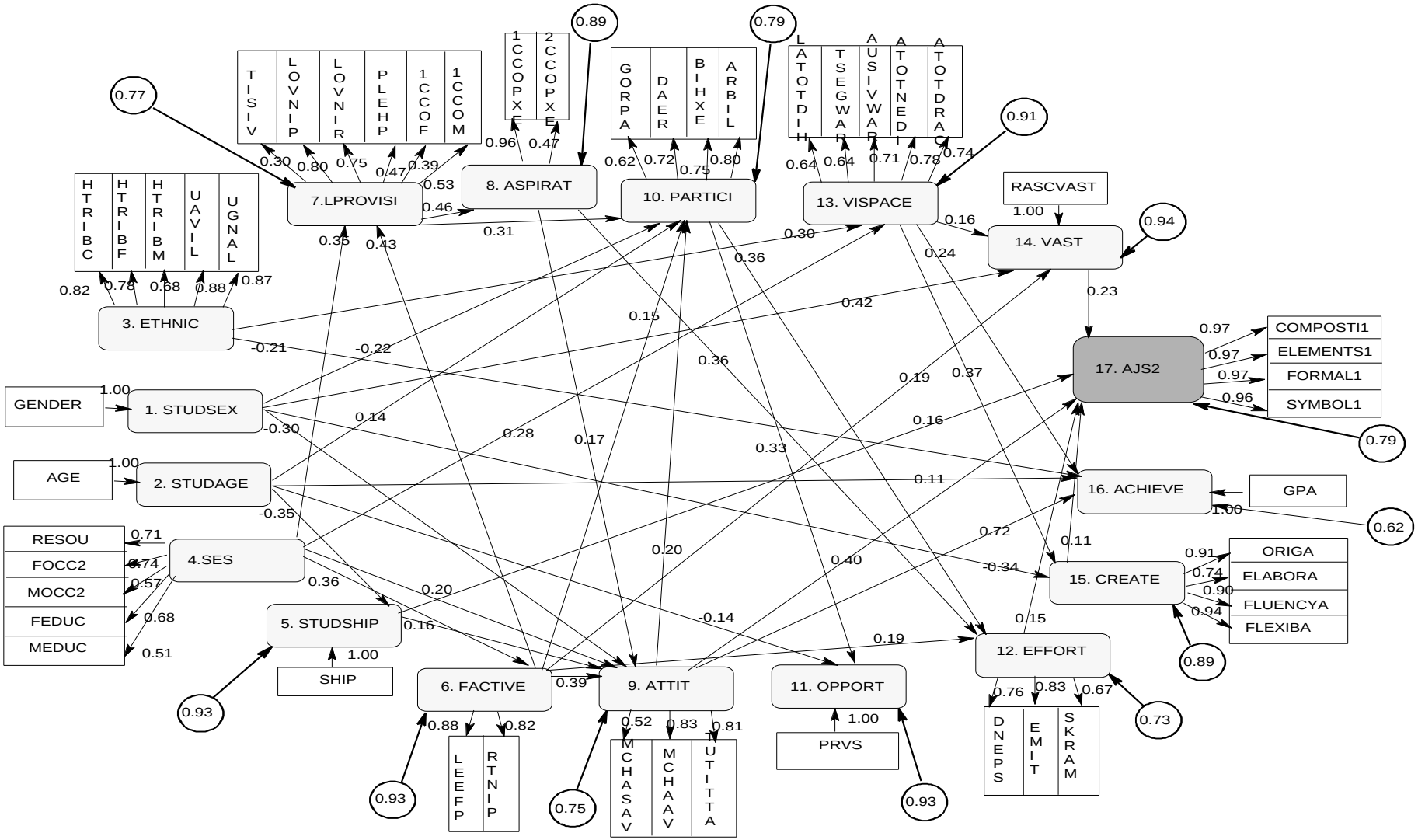


Figure 8. Grade 11 - Student level factors influencing performance on art production with AJS2 as the criterion variable.

Students' procedures for reviewing lecture notes

Tu Anh Thi Tran
School of Education, Flinders University *tuanh.tran@flinders.edu.au*

Michael Lawson
School of Education, Flinders University *mike.lawson@flinders.edu.au*

Teacher education students were questioned about the purposes they established for use of lecture notes and their use of lecture notes was observed following the presentation of a short lecture. The students also completed a test on the lecture material. The purposes students established for their lecture notes covered a wide range of study and practical tasks. Notes were reported to be most commonly used for essay writing, which was the most common form of assessment for these students. The most frequent strategy students used for reviewing of lecture notes was some form of repetition. Less common was the use of complex elaborative procedures that involved generation of links among different components of the lecture notes, or between the new material and the students' existing knowledge. Concerns are raised about the state of the students' metacognitive knowledge about lecture-note review and about the need to include such knowledge in university courses.

notetaking; lectures; lecture review; metacognitive knowledge

INTRODUCTION

Lectures remain a commonly used mode of instruction in higher education and at these lectures many of the students attending take notes on the content of the lectures. What then do the students do with these notes? It was this question that was the central focus of this project. We set out to find out how students who were studying in the field of education used the notes that they took in lectures and examined how the range of different uses of lecture notes varied in degree of elaboration. Our concern with elaboration arose from our view that the effectiveness of elaboration of information during study has a powerful influence on the storage and later use of that information.

Lectures are seen to be an efficient and flexible form of teaching that can inform students and arouse their interest in a given topic (e.g., Barbetta & Skaruppa, 1995, Barry, 1995). In these lectures lecturers expect students to take notes and students expect to take notes in lectures. Despite possibly different reasons for note-taking, most students take notes from almost every lecture (Carrier, Williams & Dalgaard, 1988; Hartley & Davies, 1978; Palmatier & Bennet, 1974). The two most common reasons for note-taking proffered by students are the usefulness of note-taking for learning and social pressure.

Hartley and Davies (1978) suggested that social pressure to take notes, from both lecturers and other students, could be associated with the need to provide "evidence of effort invested". They also noted that American students valued note-taking far more than English students and reported more social pressure than English students. It can be expected that, for students in some countries where the lecture is the most important method of teaching in higher education and notes taken from lectures are the most important resources for learning, the degree of social pressure would be very high.

There is substantial research support for the view that note-taking assists students' learning (e.g Annis & Davis, 1975; Crawford, 1925; DiVesta & Gray, 1972; Barnett, DiVesta & Rogozinski, 1981; Kardash & Kroeker, 1989). The effect on learning is argued to emerge during both encoding and storage stages of processing. The encoding effect is proposed to result from the process of attending to and recording important details of the lecture material. The storage effect is argued to be the result of the reviewing of notes and the combined effect of encoding and storage processing has been shown to be more beneficial than encoding on its own (Hartley, 1983; Kiewra, 1985d; Kiewra, Dubois, Christian, McShane, Meyerhoffer & Roskelley, 1991; Kiewra, Benton, Kina, Risch & Christensen, 1995; Peper & Mayer, 1978, 1986)

We also know that, in general terms, it is a good idea for students to review their notes. Effective review of notes taken in lectures will usually provide students with an advantage when they later undertake a related academic task such as a test. In many cases students will have access to both their own notes and those from the lecturer. Kiewra and his colleagues (Kiewra, 1983, 1985a, 1985b, 1985c; Kiewra & Dubois, 1988) found that reviewing the lecturer's notes had a greater effect on learning than reviewing students' own notes, although achieved the best performance when they reviewed both the lecturer's notes and their own notes (Kiewra, 1985a).

The form of the review used by the student is important. Peper and Mayer (1978, 1986) argued that the generative nature of processing of lecture material influences subsequent use of the lecture material. Other researchers have shown that the writing of summaries or the posing and answering of questions about the lecture material have powerful effects on subsequent use of lecture material in tests and other academic tasks (Hadwin, Kirby & Woodhouse, 1999; King, 1992). These sets of findings fit within the broad constructivist, self-regulatory perspective on student learning. The active, strategic organisation of lecture material by the student is argued to result in powerful knowledge representations that can be accessed in later problem solving. In reviewing their notes students have the opportunity to generate connections between knowledge components, either connections with prior knowledge, or with subsequent study material, or among parts of the lecture material. (Kiewra, Dubois, Christian, McShane, Meyerhoffer & Roskelley, 1991; King, 1992; Schunk & Ertmer, 1999; Winne 1995).

This is the theory that is available for use by lecturers in designing and delivering their lectures and for use by students in taking and reviewing their lecture notes. The broad objective of the current study was to see to what extent this public theory of note-taking and subsequent use of notes is also the private theory of a group of education students.

Van Meter, Yokoi & Pressley (1994, p.333) noted that "At present, we know very little about how students process their notes when they review." However, researchers have gathered some information from students about note-taking. Hartley and Davies(1978) and Van Meter et al.(1994) reported that students saw the usefulness of note-taking in maintaining their attention to the lecture. Students in the study of notetaking by Pressley, Van Etten, Yokoi, Freebern, and Van Meter. (1998) reported that the taking of notes was important for their test preparation and homework. In this research students also noted that note-taking activity facilitates comprehension and organisation of the lecture, either by helping them to recognise the structure of lecture material during the lecture or by helping them to restructure it after the lecture. Students also regarded note-taking as providing them with material for revision and referencing (Hartley & Davies, 1978; Van Meter et al., 1994).

The studies by Van Meter et al. (1994) and Pressley at al. (1998) were principally studies of note-taking practices. Students distinguished between lectures that were easy-to-note and those that were made difficult to note by such factors as lecturing style, pace of presentation, quality of explanation and organisation. Students also reported varying their note-taking styles according to their prior knowledge, lecturer style, experience

The current study is focussed on students' use of their lecture notes and was designed to provide further knowledge about the processes students use to review their notes. In seeking student explanations for their views we also aimed to gather information about what Van Meter et al, (1994) describe as the students' theory of note-taking, or their theory of use of notes. Specifically the current study was designed to provide information on the following issues:

1. The purposes established by students for lecture notes.

2. The use made of notes and the reasons for this use.

3. The strategies seen to be most effective for reviewing notes, and why these strategies were regarded as effective.

In seeking this information were aiming to provide information about the degree of match between the students' theories of note-taking and use of notes and the account generated in the self-regulatory approach to learning in higher education.

METHOD

Participants

The study reported here formed part of a larger study. In the initial part of the study there were 55 students in the final years of a 4 year education program that attracted both undergraduate and postgraduate entry. Twenty-nine students were in the third year of their program. For the final parts of the study a subset of 18 of these students provided data. This group of students were the ones from the larger group who had not received any specific strategy training during the larger study.

Procedure

Students took part in four different activities. First they completed a questionnaire which asked them to describe what they did with their notes after a lecture. They then watched an 18 minute videotaped lecture on theories of forgetting and took notes in any way they wished, knowing that there would be a short-answer recall test at the end of the session. Following completion of the lecture the smaller subgroup of 18 students read a short article on hypnosis. This was a filler activity designed to provide a short period of alternative activity. After this the students were asked to review their lecture notes and also had access to a lecture outline.

Students then completed a second questionnaire that focussed on the procedures they used for reviewing notes. Questionnaire 2 consisted of two open questions. In the first question, students were asked to describe all the review procedures they used when they reviewed their notes from the lecture on forgetting. The second question asked them to indicate the review procedure they found most effective for their learning and to explain the way this effective procedure would help their learning. A recall test on lecture content was completed after this.

In Questionnaire 3 students were asked to select from a list the three lecture-note review strategies those strategies they regarded as most useful for learning, and the three they regarded as least useful. They were also asked to explain the reasons for these selections.

Coding of responses

The responses to all questionnaires were coded using the same broad procedure. Codes were generated from an analysis of the constructs of elaboration and complexity used in the literature of instructional psychology (e.g. Anderson, 2000; Pressley and McCormick, 1995). The focus in this case was to generate categories that differed in degree of elaboration or complexity. Code definitions were first discussed and definitions agreed upon by two raters. These definitions were then used in a first coding of questionnaire responses by one of the raters. Any difficulties in definition or coding of specific responses were then discussed by

both raters and modifications to the coding scheme agreed upon. A second coding was then carried out by one rater, with a reliability check being done on a small sample by the second rater. The final rating was not undertaken until both raters resolved differences in rating of specific cases.

Analysis of responses in Questionnaire 1

Responses to Questionnaire 1 were separated into those that indicated a *purpose* for use of lecture notes and those that described a *strategy* that the student used with the notes. Three major categories of purposes were reported; namely purposes related to specific academic tasks, purposes for general learning, and purposes related to career. More specific groupings within these three categories were also identified. In the categorising of statements about strategies for using lecture notes, the responses were grouped into nine different types of strategy; namely filing, seeking of reference, repetition, highlighting, sounding, simple elaboration, complex elaboration, selection and monitoring. A definition and an example of each strategy is shown in Appendix 1.

Analysis of responses in Questionnaire 2

Questionnaire 2 collected information about students' actual activities engaged in while reviewing their notes from the lecture. Therefore, information obtained from Questionnaire 2 related to a specific situation rather than to the general use of lecture notes considered in Questionnaire 1. Three parts of students' answers were identified; the strategies actually used for reviewing lecture notes, the strategies nominated as most effective strategies for review, and the explanations given for effectiveness of these strategies.

The strategies used for reviewing notes made during the lecture on forgetting were coded using the approach developed for analyzing responses from Questionnaire 1. A strategy was coded as effective when it was reported by students as effective for review, or when it could be inferred as an effective activity from the answer. In their answers, some students wrote explicitly that the strategy was "the most effective" or "the most useful" for them. In other cases, effective strategies could be identified from a "cause and effect" structure in students' descriptions. If the activity was considered as "cause", it was coded as an effective strategy. On the contrary, if the activity was considered as "effect", it was coded as a reason for the effectiveness of review strategies. For example, in the answer "*Forming a picture. If I have a picture in my head I can relate it to specific circumstances*" "forming picture" and "have a picture" were seen as a "cause" and "can relate it to specific circumstances" as an "effect". Therefore, the former was included in the group of effective strategies and was coded as *visualising*, and the later was coded as a reason for effectiveness, in this case *linking*. The labels, definitions and examples of each category of effective strategies are included in Appendix 2. The analysis of students' reasons for the effectiveness of strategies will be introduced with the analysis of the Questionnaire 3 below.

Ratings of complexity

In both Questionnaire 1 and Questionnaire 2 strategies were rated as involving *simple* elaboration or clarification when they did not make any change to the content of lecture notes. These strategies were for example, repetition, sounding or triggering memory. Simple strategies received the complexity score of 1. Strategy descriptions that involved transformation of content in the lecture notes were divided into two categories depending on the degree of change that could result from use of the strategy. In the *less complex* category, strategies did not involve either change to connections within concepts, or change to the relationship between new information and student's prior knowledge. Strategies in this less complex strategies were for example, expanding notes, visualising and making simple interpretations or clarifications. These strategies were given a score of 2 for complexity.

Any strategy that involved a change in connections within concepts of the lecture and/or between new information and student's existing knowledge was placed into *more complex* strategy group. These strategies were, for example, complex clarification, organisation, paraphrasing, summarising, selecting, linking, or explaining. All of these strategies were given a score of 3 for complexity.

Analysis of responses related to reasons for effectiveness and usefulness of strategies

Reasons for the usefulness of review strategies were obtained from students' explanations for their choice of a useful strategy in Questionnaire 3 using the same process as that used for the reasons for effective strategies. Reasons for the effectiveness and the usefulness of review strategies were coded using the categories and definitions presented in Appendix 1. The procedure used to rate was the same as that used for rating the strategy descriptions given in this questionnaire as described above. The complexity score of each student was the total score for all reasons reported.

The categorising of explanations why specific review strategies were regarded by students as being *not* useful for study purposes followed the logic used in categorising of responses from the other questionnaires. Categories were established to recognise different degrees of transformation of note content, using the simple, less-complex, and more-complex categories used previously. The remaining categories emerging in this analysis were of a different nature to those that had been identified in previous analyses, being concerned with social and efficiency concerns. The categories used in this analysis are shown in Appendix 4.

RESULTS

Purposes described by students for using lecture notes

The focus of interest here is the purposes that students saw for using notes after lectures. While other research has been concerned with students' purposes for note-taking in general (Hartley & Davies, 1978; Isaacs, 1994; Van Meter et al., 1994), this study focused more specifically on the purposes that students suggest for using notes after lectures.

Three main groups of purposes of using notes were identified, namely "specific purposes for learning", "purposes related to career" and "purposes for general learning". Details of the type of purpose and the percentage of each reported purpose in relation to the total number of reported purpose answers are presented in Table 1.

Over half of the purposes identified by students were those concerned with specific academic tasks, such as essay writing and exam preparation. A smaller percentage of responses indicated that notes could be used in preparation of tutorials and seminars. One third of the responses focused on more general learning purposes such as future study or understanding and a small number of expressed purposes (6%) were related to students' practical teaching.

The use of lecture notes for writing essays was the most common single purpose seen for lecture notes, representing 32% of responses. Lecture notes could be used as a guide for essay writing in planning, organising or as a checklist which helped students "*make sure I have not forgotten to cover something important*". Lecture notes could also provide students with information, key points, and statements that could be quoted in an essay. Lecture notes were seen to be useful for general preparation for exams and as materials for revision before the exams. This result is consistent with data obtained from the Van Meter et al.'s (1994) study. Lecture notes could also help students to "*prepare for talks/ presentations"*, to "*respond to questions in workshops"* or to "*recheck details/ information"*. The low frequency of response about using notes for tutorials and seminar preparation might indicate that many students do not draw explicit links between their lecture and tutorial/seminar classes.

Table 1. Purposes for using lecture notes

Specific academic tasks (56%)	Essay writing (32%)	General use for writing essays (9%)	
		Reference for writing (8%)	
		As guide for writing (6%)	General guidance (3.5%)
			For process of writing (1.5%)
			As a checklist (1%)
		Information provision (6%)	Key ideas (4.5%)
			General information (1.5%)
		Quote (3%)	
	Exam preparation (15%)	Revision for exams (9%)	
		General preparation (6%)	
	Tutorials and seminars (9%)		
General learning purposes (38%)	Future study (21.5%)	References (12.5%)	
		Next lectures (3.5%)	
		Other subject (3.5%)	
		Further study (2%)	
	Understanding (5.5%)	Clarification (3.5%)	
		Overview of a topic (1%)	
		General understanding (1%)	
	Memory jogger (4.5%)		
	Monitoring (3.5%)		
	Revision (3%)		
Purposes related to career (6%)	Teaching practice (3%)		
	Working in school (3%)		

The high frequency of reported use of notes for essay writing may well be associated with the high frequency of using essays as a method of assessment in the department where the students are studying. The students in this study were enrolled in a faculty where most of their assessment, though not all, was based on a combination of written assignments and examinations.

Consistent with the findings of Hartley and Davies' (1978) and Isaacs (1994), students in this study reported the use of notes for further study. Future study could be to *"use sources cited as a basis for further reading", "preparing for next lectures", "use in other subjects"* or *"as a basis for further development of lecture material"*. Lecture notes could be used to help students in understanding and remembering lecture materials and for monitoring learning, suggested by views such as "*Notes are used to clarify any misunderstanding that I might have about a subject*" and "*as a prompt to recall information learned in lecture"*.

The students' field of study obviously plays an important role in the purposes of using lecture notes. As the students in this study were student teachers, there were a small number of reports that lecture notes were used for teaching practice or for working in school. These referred to instances where they "*use notes as a basis or for ideas when planning lessons for teaching practices*" or use notes "*when writing curriculum for a class*".

In summary, most students did report that notes were used after the lecture. The emphasis in the purposes reported may reflect the interaction between the method of learning and assessment in the students' degree program. Although the findings generated from Questionnaire 1 show a similar general pattern to those obtained from previous studies, they

do provide more specific information about students' purposes for using notes after lectures. The results show that the student group made some reasonably fine discriminations among uses for lecture notes, though most of the categories of use were still rather general. It was the nature of the categories that was the focus of the next analysis.

Types of strategies reported by students for review of their lecture notes

The interest of this part is the nature of the procedures adopted by students for using notes during study. The analysis of the students' reports on these procedures was designed to see to what extent the procedures involved substantial transformation of lecture notes that might is associated with more elaborate, or deep, learning (Pressley and McCormick, 1995). The types of procedures and the number of students reporting use of each are presented in Table 2. The definitions used in making decisions about grouping of these strategies are shown in Appendix 1.

Table 2. Frequency of use of different type of review strategy

Type of procedure	**Number of students reporting use (n=55)**
Filing	12
Reference	24
Repetition	42
Highlighting	5
Selection	6
Rhyming	2
Monitoring	1
Simple elaboration	10
Complex elaboration	16

Of the 55 students, 12 reported that they usually filed lecture notes for future use. This strategy was reported as simple filing, such *as "filing away without doing any thing else."* Almost a half of the students reported that they used their lecture notes as a source of references for essay writing, for seminar presentations, or for tutorials. It also indicates that students looked to lectures as an important source of information for further study. We are not able to say from this study how this use for referencing compares with the use of other sources such as set readings or library research. The findings here are consistent with results obtained from Hartley and Davies (1978).

Repetition was by far the most common procedure reported by students. Forty-two of the 55 students reported activities such as rereading or rewriting, a pattern consistent with the findings in Van Meter et al.'s (1994) study, and those of Kardash and Kroeker (1989) who reported that rereading was the most popular review strategy that was reported by students. The emphasis on repetition in reviewing lecture notes is consistent with a wide use of this type of strategies in second-language vocabulary learning reported by Lawson and Hogben (1996).

Small numbers of students reported that they selected particular parts of notes for attention through use of procedures such as highlighting or underlining to identify "*major areas*", or "*key points*". Two students reported use of rhymes to help them in memorising and only one student reported using notes to monitor understanding .

Twenty five students reported some form of elaboration of lecture note material. Simple elaboration involved activities such as "*expanding lecture notes*", "*visualising*", "*ordering*" or "*making simple clarifications*", which made only simple changes or additions to the lecture notes. The complex elaboration procedures, such as " *outlining*", "*evaluation*", "*concept mapping*", "*linkin*g", "*paraphrasing*" and "*summarising*" would have greater potential to result in the generation of more connections among ideas in the lecture material, or between

that material and students' prior knowledge. This latter activity would be predicted to be of more use for subsequent complex problem solving.

The pattern of results in Table 2 suggests most of these final year students reported more frequent use of such simple strategies as repetition than of more complex strategies such as elaboration. Although elaboration is believed to be a more effective learning strategy for longer term retention of information than repetition, many of the students in this study did not report use of the more complex procedures when reviewing their notes. Spontaneous use of elaborative strategies has not been reported by many of the students in either the current study or in other research (e.g., Pressley, Wood, Woloshyn, Martin, King & Menke, 1992). As noted earlier it seems likely that the strategies adopted for use of notes will be influenced by the types of learning and assessment activities established for students by their lecturers. This should be an area for further investigation.

Number of strategies reported by each student

We were also interested in the number of types of strategies mentioned by each student. The number of strategies mentioned by each student ranged from none to six. Two students did not report use of any strategy. Most reported use of two strategies (24), with 12 students reporting use of a single strategy and 13 reporting use of 3 strategies. Four students reported use of four or more strategies. When the complexity of strategies reported in Table 2 was rated, there was a strong correlation between the number of strategies reported and strategy complexity ($r = .93$, $p < 0.01$). This result suggests that the greater number of strategies used by students, the more likely they were to include complex strategies.

Review strategies used with notes from the lecture on forgetting

The focus in this part of the analysis is on what strategies students actually used to review lecture notes from a specific lecture on the topic of forgetting. These results came from the answers of the subgroup of 18 students who did not receive strategy training in the larger study. A point of interest with this data was that the complexity of the reported procedures could be related to the scores students obtained in the test following the lecture on forgetting. The data related to this question is shown in Table 3.

When compared with the pattern observed in the responses in Table 2, we again see that repetition was the most frequently reported strategy. However, in the case of this specific lecture, a higher percentage of students reported use of simple and complex elaboration strategies than was reported for the general situation in Questionnaire 1. In looking for reasons why there was greater use of elaboration in this part of the study we suggest that the availability of a detailed lecture outline may have encouraged more generative activity in students, even though this outline was organised in a linear manner. The work of Kiewra (1983) suggests that outline format can be a strong influence on the type of transformation of lecture material used by students. The complexity of the reported strategies was again rated and there was a moderate correlation between the test score obtained by students and the complexity of strategies used for reviewing lecture notes, ($r=0.51$, $p < 0.05$.) The more complex the strategies students used for reviewing of lecture notes, the better they performed on the test.

Strategies considered as effective

In the final section of the analysis of information gathered in this study we were interested in the students' reasoning about the strategies they employed when using lecture notes. This information might be seen as giving us a glimpse of the students *theories* about use of their lecture notes. The question related to this purpose asked students to identify the strategies they regarded as being most effective for reviewing of the notes they took during the lecture on forgetting. The results are shown in Table 4.

Table 3. Frequency of use of different type of review procedure for lecture on forgetting

Categories of strategies	Number of students reporting use (n=18)
Repetition	11
Highlighting	3
Selection	5
Rhyming	4
Expanding notes	5
Simple elaboration	7
Complex elaboration	7

The responses follow a similar pattern to that observed in the students' reports of the strategies they used during this lecture and support the usefulness of those reports. Simple strategies regarded as effective involved use of the lecture handout or of the students' own notes for activities such as self-testing, reading aloud, highlighting and repetition. Students also reported that simple elaboration strategies such as expanding notes, adding keywords or examples, visualising, interpretation and simple clarification were effective. A smaller percentage of the students identified more complex elaboration as being effective. These activities included activities such as organisation, linking of ideas and the development of explanations for points noted from the lecture.

Table 4. Frequency of effective strategies reports

Effective strategies	Number of students (n= 18)
Repetition	10
Other simple strategies	4
Simple elaboration	7
Complex elaboration	5

Reasons given for why review strategies were effective or useful

We were interested in the reasons students provided for the effectiveness and usefulness of review strategies. The results for this analysis came from Questionnaires 2 and 3.

The students' reasoning about strategy effectiveness and usefulness was assumed to be based upon their understanding of the procedures that would result in good quality learning. It is this understanding that we refer to as comprising the students' theories of learning.

Reasons for effectiveness or usefulness were categorised into three main groups, namely, simple reasons, less complex reasons and more complex reasons. Categories of reasons and the number of students giving each reason are shown in Table 5. The number in parenthesis in each cell is the total number of responses in that cell.

Strategies such as memory triggering and repetition were seen to be effective because they addressed both cognitive and affective needs of students. These procedures could provide cues for recall and focus attention, both useful cognitive actions. They could also increase the level of confidence of the student and assist with referencing in subsequent work. However, these *simple reasons* did not involve any change of the learning materials.

Seven types of *less complex reasons* were identified, all of which related to some simple changes in the learning materials. In this group of reasons, an undifferentiated understanding was the most common reason given for the usefulness of a strategy. Moreover, as simple clarification, simple personalising and interpretation could result in a better understanding, they supported the students' strong need for understanding. What is not apparent in this group of

reasons is a concern for generating interaction among the different components of a set of notes.

Table 5. Frequency of reason categories for strategy effectiveness (Q2) and strategy usefulness (Q3)

Reasons	**Reasons**	**Effectiveness(Q2)**	**Usefulness(Q3)**
	Repetition	4	2
	Memory trigger	4	1
	Attention	2	1
Simple reasons	Sounding	1	1
	Confident	1	
	Referencing		1
	More time for learning		1
	(Subtotal)	(12)	(7)
	Understanding	1	5
	Simple clarification	1	1
Less complex	Visualising	3	2
reasons	Simplifying		1
	Simple personalising		1
	Interpretation		1
	(Subtotal)	(5)	(11)
	Selecting		1
	Paraphrasing	1	1
More complex	Linking	3	4
reasons	Organisation	2	2
	Summarising		1
	Explaining		1
	Placing in context		1
	(Subtotal)	(6)	(10)

This generative, or interactional, activity was the focus of the *more complex reason* group. The reasons advanced in this group focused more on bringing about interactions between students' prior knowledge and new lecture information, or interaction within components of the lecture materials. As was the case in earlier analysis, there was less frequent discussion of the more complex elaborative strategies than of simpler features of learning.

There was a difference in the pattern of response to the effectiveness and usefulness questions. Most prominent is the increase in the frequency of nomination of elaborative activities as being useful for learning. It is important to remember here that in making the judgements about usefulness students were making choices among a list of provided strategies, rather than generating their own justifications. The list of options seems to have reminded them of alternatives that were not as readily accessed when they were responding without such reminders.

A comparison of the pattern of responses shown in Tables 2 and 3, and those in Table 5, raises the possibility of a difference between students' theories and their practice. In the earlier tables we saw that most of the students reported use of repetition, either in general or after the lecture on forgetting in this study. A lower percentage of students focussed on use of elaboration in those reports. Yet, in Table 5 the frequency of more complex reasons is quite similar to that of similar reasons, especially in the responses to Questionnaire 3. Within the limits of the current study we cannot resolve this possible inconsistency between students' practice and their reasoning about review strategies and it does seem to be a fruitful area for further investigation.

We rated the complexity of the students' reasons for usefulness of strategies and correlated these ratings with their scores on the final test There was a significant moderate positive correlation between the final total test score and the complexity score of reasons for usefulness ($r= 0.56$, $p< 0.05$).

Why do students regard strategies as not useful for learning?

The final set of information coming from Questionnaire 3 referred to students' reasons for regarding review strategies as not being useful for their learning. As was the case for their decisions about usefulness of a strategy, students were asked to explain why they regarded a nominated strategy as not being useful. Categories of reasons for why review strategies were not seen to be useful and the number of students answering in each category are shown in Table 6. Not all students were able to, or chose to, respond to this question.

As can be seen from Table 6, the reasons given here were varied and mostly of low frequency. However, they do provide some different information about students' reasoning than was available from responses to other questions. Two responses referred to the social context of study that had not been mentioned previously. One student stated that "*I think I am a social learner, therefore doing all the thinking in my head doesn't allow me to learn as effectively as talking about it out aloud does*".

Efficiency factors were is the largest group of reasons for the non-usefulness of review strategies. A review strategy was not useful if a student's time and effort were used inappropriately because the strategy " *requires too much effort and time*". In addition to time and effort, students were concerned about the amount of to-be-learned information. Review strategies, although they were proved effective in learning, were still not useful if they increased the amount of the to-be-learned information. For example, one student found review by associating each main idea with a letter or object and memorising the list of letters/ objects was not useful because "*by adding symbols etc. it adds to more information that you have to remember*". In a similar way one student rejected a review strategy because it increased the difficulty of the review process. Adding symbols when reviewing notes was not useful because "*symbols can ... become confusing, especially when under pressure in an exam*".

Only two of the responses referred explicitly to reasons associated with complex processing of the lecture material. One student stated that: "*I find rote learning difficult because it removes the context from the information. It strips the information of any triggers to assist in recalling at a later date*."

Table 6. Frequency of category use given by students for why strategies are not useful

Reasons		Number of answers
Social reasons	Lack of social context	2
	Time consuming	1
	Effort consuming	1
Efficiency reasons	More information to learn	2
	Not relevant	1
	Redundancy	1
Difficulty reasons	Confusing	1
Simple reasons	Forgetting the letters	1
	Misplace material	1
	Only repetition	1
Less complex reasons	No understanding	1
More complex reasons	No link	1
	Removes context	1

One implication of this admittedly small set of responses to this final question is that they serve as a reminder that strategy knowledge is not just cognitive - it also has an affective component. Students will not use what might be a powerful strategy if they do not like using it.

CONCLUSION

Pressley at al. (1998) describe this area of research appropriately by referring to it as "*the metacognition of college studentship*." The knowledge about notetaking that students take with them to lectures and to the use of their notes after lectures is functionally important. It is this knowledge that guides their use of the notes they take and it is the outcome of use of this knowledge that will influence their subsequent problem solving on the topic of the lecture. The results of our analysis add to our knowledge about the 'theories' that students have constructed in this area of university studentship.

Students report use of notes for a wide variety of tasks in both their subsequent study activities and in their practical work as teachers. In the group of students involved in this project lecture notes provided a major resource for support of the essay writing that formed the bulk of their assessment tasks.

Of more interest in this study was the way in which students used their notes and it is the results of this analysis that is of most relevance for the research on the metacognition of studentship. Repetition of material in notes was the most frequently reported strategy for both lectures in general and for the short lecture that formed part of this study. Less frequently reported was the use of strategies that involved complex elaboration of the lecture material. We regard this pattern of results as indicating an area of concern for lecturers and students. Of course the degree of concern that is appropriate here is dependent on the requirements established for the use of lecture material in the assessment tasks that form part of the student's course or program.

Here, let us focus on assessment tasks that require a deep understanding of the lecture material. For these tasks we can reasonably argue that the student will need to be engaging in complex elaboration of much of the material included in lecture notes. This will be the case wherever the assessment task requires some extension of what has been learned in lectures, or when it involves some novel perspective needs to be adopted on the lecture topic, or where far-transfer problems must be solved. In all of these cases the student would be predicted to benefit from complex transformation of the lecture notes. The correlational analyses in this study relating review strategy and test score support this view.

An issue of concern for lecturers in our findings is the range of knowledge that students have about effective uses of lecture notes. In our data, those students who used more strategies also were more inclined to know about and use more complex elaboration strategies, which was positively associated with higher test score. Yet not all our students reported that they typically used complex elaboration strategies or reported use of them when they would have been useful in this study. This raises a question about whether students know as much as they need to about such strategies. If our answer to this question is in the negative we then need to consider whether we should be teaching students such strategies as part of our courses. We are inclined to answer in the positive to this last question.

REFERENCES

Anderson, J. R. (2000) *Cognitive Psychology and its Implications* (5th ed.). New York: W. H. Freeman & Co.

Annis, L & Davis, J. K. (1975) The effect of encoding and an external memory device on note-taking. *Journal of Experimental Education*, 44, 44-46.

Barbetta, M. P. & Skaruppa, L. C. (1995) Looking for a way to improve your behaviour analysis lectures? Try guided notes. *The Behavior Analyst*, 18, 155-160.

Barnett, E. J., Di Vesta, J. F. & Rogozinski, T. J. (1981) What is learned in note-taking? *Journal of Educational Psychology*, 73 , 181-192.

Barry, K. (1995) Lecturing, explaining and small-group strategies. In Maltby, F., Gage, N. L. & Berliner, D. C. (Eds.) *Educational Psychology: an Australian and New Zealand Perspective*, John Wiley & Sons.

Carrier, C. A., Williams, M. D., & Dalgaad, B. R. (1988) College students' perceptions of note-taking and their relationship to selected learner characteristics and course achievement. *Research in Higher Education*, 28, 223-239.

Crawford, C. C. (1925) Some experimental studies of the results of college note-taking. *Journal of Educational Research*, 12, 379-386.

Di Vesta, F. J. & Gray, G. S. (1972) Listening and note taking. *Journal of Educational Psychology.* 63, 8-14.

Hadwin, A. F., Kirby, J. R., & Woodhouse, R. A. (1999) Individual differences in notetaking, summarization and learning from lectures. *Alberta Journal of Educational Research*, 45, 1-17.

Hartley, J. (1983) Note-taking research: resetting the scoreboard. *Bulletin of the British Psychological Society*, 36, 13-14.

Hartley, J & Davies, I. K. (1978) Note-taking: A critical review. *Programmed Learning and Educational Technology*, 15, 207-224.

Isaacs, G. (1994) Lecturing practices and note-taking purposes. *Studies in Higher Education*, 19, 203-216.

Kardash, C. A. M. & Kroeker, T. K. (1989) Effects of time of review and test expectancy on learning from text. *Contemporary Educational Psychology*, 14, 325-335.

Kiewra, K. A. (1983) The process of review: A levels of processing approach. *Contemporary Educational Psychology*, 8, 366-374.

Kiewra, K. A. (1985a) Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology*, 10, 378-386.

Kiewra, K. A. (1985b) Learning from a lecture: An investigation of note-taking, review and attendance at a lecture. *Human Learning*, 4, 73-77.

Kiewra, K. A. (1985c) Providing the instructor's notes: An effective addition to student notetaking. *Educational Psychologist*, 20, 33-39.

Kiewra, K. A. (1985d) Investigating note-taking and review: A depth of processing alternative. *Educational psychologist*, 20, 23-32.

Kiewra, K. A., Benton, S. L., Kina, S. I., Rich, N. & Christensen, M. (1995) Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology*, 20, 172-187.

Kiewra, K. A., & Dubois, F. N. (1988) Providing study notes: Comparison of three types of notes for review. *Journal of Educational Psychology*, 80, 595-597.

Kiewra, K. A., Dubois, F. N., Christian, D., McShane, A., Meyerhoffer, M. & Roskelley, D. (1991) Note-taking functions and Techniques. *Journal of Educational Psychology*, 83, 240-245.

King, A. (1992) Comparison of self-questioning, summarizing, and note-taking- review as strategies for learning from lectures. *American Educational Research Journal*, 29, 303-323.

King, A. (1994) Guiding knowledge construction in the classroom: Effects of teaching how to question and how to explain. *American Educational Research Journal*, 31, 338-368.

Lawson, M. J. & Hogben, D. (1996) The vocabulary-learning strategies of foreign-language students. *Language Learning*, 46, 101-135.

Palmatier, A. R. & Bennett, J. M. (1974) Note-taking habits of college students. *Journal of Reading*, 18, 215-218.

Peper, J. R. & Mayer, E. R. (1978) Note taking as a generative activity. *Journal of Educational Psychology*, 70, 514-522.

Peper, J. R. & Mayer, E. R. (1986) Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, 78, 34-38.

Pressley, M. & McCormick, C. B. (1995) *Advanced Educational Psychology for Educators, Researchers and Policymakers*. Harper Collins College Publishers, NewYork.

Pressley, M., Van Etten, S., Yokoi, L., Freebern, G., & Van Meter, P. (1998) The metacognition of college studentship: A grounded theory approach. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 347-366). Mahwah, NJ: Erlabum.

Pressley, M., Wood, E., Woloshyn, V., Martin, V., King, A. & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27, 91-109.

Schunk, D. H. & & Ertmer, P. A. (1999) Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology*, 91, 251-260.

Van Meter, V. P., Yokoi, L. & Pressley, M. (1994) College students' theory of note-taking derived from their perceptions of note-taking. *Journal of Educational Psychology*, 86, 323-338.

Winne P. H. (1995) Inherent details in self-regulated learning. *Educational Psychologist*, 30, 173-188.

APPENDIX 1:
DEFINITIONS AND EXAMPLES OF STRATEGY CATEGORIES IN QUESTIONNAIRE 1

Category	Definition	Example
Filing	Storing lecture notes	File with other lecture notes.
Reference	Getting information from lecture notes	Look up names of authors or readings mentioned by the lecturer.
Repetition	Accessing lecture notes at least once more after the lecture.	Reread, or rewriting, the notes taken from lectures.
High-lighting	Emphasising parts of lecture notes	Underline key words, ideas in lecture notes.
Sounding	Making parts of lecture notes become auditory	I make up rhymes with the notes gathered.
Simple elaboration	Changing lecture note content in simple ways, without establishing new relationships.	Expand notes into sentences/ paragraphs for easier use, reordering, or clarifying.
Complex elaboration	Changing connections within lecture notes or between prior and new knowledge.	Write up in a summary of lecture notes, or mind mapping.
Selection	Distinguishing main ideas from less important ones	Pick out most important and relevant points...
Monitoring	Monitoring understanding or learning activities	For exams I look, cover, check.

APPENDIX 2:
STRATEGIES REGARDED AS EFFECTIVE FOR REVIEW

Strategy	Definition	Example
Repetition	Accessing lecture notes at least once more	Writing the information out several times helps me learn effectively.
Triggering memory	Helping students to "trigger" or "jog" their memory	Using notes to trigger memory of actual event in video.
Sounding	Making lecture content become auditory.	Saying it over in my head helps.
Expanding notes	Adding information to lecture notes	Writing the additional notes from the sheet and adding these to my notes was most useful…
Visualising	Making lecture notes visual or forming picture	Visualisation helps me bring relevance and a humanistic approach to my study.
Interpretation	Adding interpretation to lecture notes	I think that adding my own interpretation of what was in lectures helps me learn most effectively
Simple clarification	Making clarification in simple ways, without changing the structure of lecture notes	Clarification is essential in avoiding misconception.
Complex clarification	Making clarification which related to the change in the structure of lecture notes	Maybe the clarify one. (Try to give it more structure)
Paraphrasing	Restating, rephrasing lecture notes in own words	... put the words in my own way- paraphrasing help me to understand what is being said.
Selecting	Distinguishing main points from less important ones	Repeating main points to myself. (…) Get the important stuff memorised first.
Organisation	Changing the structure of the lecture notes.	Organisation of the information was important for me to help in my retention.
Linking	Relating, connecting points in lecture notes or between new and old knowledge	Trying to develop understanding and sometimes relate it to something.
Explaining	Giving explanations to concepts in lecture notes	I find explaining the concept to people helps me learn most effectively.

APPENDIX 3:
REASONS WHY REVIEW STRATEGIES WERE REGARDED AS EFFECTIVE OR USEFUL

Category	Definition	Example
Repetition	Helps students access lecture notes at least once more after lectures	By rewriting my notes, I am going over the information again...
Memory trigger	Helps students to "trigger" or "jog" their memory	Visualisation helps me trigger memories and ideas on the subject
Attention	Helps students in paying attention to some parts of lecture notes	Rewriting notes causes me to focus on every points as I go through
Sounding	Helps students "hear" the lecture content better	I am good at remembering rhythmical patterns.
Confident	Using strategy makes students feel confident	If I feel sure that my interpretation is accurate I am more likely to feel confident in recall.
Referencing	Helps students to get information from lecture notes	... being able to find information in notes has proven adequate.
More time for learning	Helps students to have more time for leaning	Rewriting allows extra time to go over material slowly.
Understanding	Making students have a better understanding of lecture content	Often I rewrite notes several times and that also helps me to understand.
Simple clarification	Helping students clarify concepts	Talking with other people clarifies any inconsistencies...
Visualising	Helping students "see" the lecture material or build a mental picture of the notes	A mental picture of the notes is built up-ideas underlined or highlighted are visualised when recalled for test or assignment.
Simplifying	Helping students make complicated information more simple.	Examples can make complicated information more simple...
Simple personalising	Helping students make information become more suitable to themselves	I feel is important in that learning is most effective when it has personal relevance...
Interpretation	Helping students to interpret lecture notes	Talking with other people allows me to explain my interpretation.
Selecting	Helping students distinguish main points from less important ones	Repeating main points to myself. Get the important stuff memorised first
Paraphrasing	Helping students restate lecture notes in their words or in their ways.	Writing the additional notes…was most useful because I was able to put the words in my own way.
Linking	Helping students to "link" or "relate" ideas in lecture notes to each other	It allows me to put the information into a way I feel makes me remember by linking words or ideas...
Organisation	Helping in organising lecture notes or giving a structure to them	I go over my notes and try to give them more structure to put information into groups.
Summarising	Helps students to summarise the lecture notes	This is useful because images... can often summarize a lot of information...
Explaining	Helping students to "explain" the meaning of the lecture notes	Talking with other people... allows me to explain my interpretation of the information.
Placing in context	Helping students to place information into the context	My mental images place me back in the context of when I learnt the material...

APPENDIX 4:
REASONS WHY REVIEW STRATEGIES WERE NOT REGARDED AS USEFUL

Social reasons	Lack of social context
	Time consuming
	Effort consuming
Efficiency reasons	More information to learn
	Not relevant
	Redundancy
Difficulty reasons	Confusing
Simple reasons	Forgetting the letters
	Misplace
	Only repetition
Less complex reasons	No understanding
More complex reasons	No link
	Removes context

Interviews with teachers and learners

Helen Askell-Williams
Flinders University, School of Education *helen.williams@flinders.edu.au*

Over the past few months I have been interviewing teachers and learners about their understandings about teaching and learning. This paper follows from those interviews and begins to discuss what teachers and learners told me. From an individual perspective, interviewees discussed their motivations, ability assessments and strategies for teaching and learning. From a contextual perspective, interviewees spoke about social interactions, the relative ease of learning in practice and the enhanced learning to be gained from combining theory, practice and reflection. In this paper I also discuss comparisons and contrasts between interviewees' understandings and current recommendations for best practice found in extant teaching and learning literature.

Teachers, Learners, Understandings, Interviews, NUD*IST

INTRODUCTION

My review of philosophical, psychological and classroom practice literature suggests that teachers' and learners' understandings, intentions and plans for action are key determinants of teaching and learning behaviours (for example, see Bereiter & Scardamalia, 1989; Hirst, 1971; Kerr, 1981). The literature also reports many examples of *mis*understandings between teachers and learners (for example, see Cusworth, 1995; Hogan, 1999; Tasker & Freyberg, 1985; Winne & Marx, 1982). Understandings and misunderstandings seem worthy of further investigation.

The literature also highlights many advances that have been made in teaching and learning theory and pedagogy over the last thirty or so years (for an overview see Bransford, Brown, & Cocking, 1999). This added another dimension to my initial concern for whether teachers and learners share similar understandings: that is, whether those understandings reflect current theory and recommendations for best practice as reported by researchers and reflective practitioners.

BACKGROUND

This paper is founded upon three areas of theory and research, reflecting a growing awareness in the research community of the logic, utility and fruitfulness of *consilience*, "a 'jumping together' of knowledge by the linking of facts and fact-based theory across disciplines to create a common groundwork of explanation" (Wilson, 1998 p. 8): (for example Anderson, Greeno, Reder, & Simon, 2000; Cobb, 1994; Entwistle, Skinner, Entwistle, & Orr, 2000; Newsome III, 2000). These three foundations are

1. A philosophy of teaching and learning which prescribes that intentions predicate plans and actions, and which draws a distinction between schooling and teaching (Bereiter & Scardamalia, 1989; Kerr, 1981).

2. Social-cognitive theory, which proposes a transaction between cognition, behaviour and environment, in particular the motivational and self-regulatory aspects of human behaviour (Bandura, 1997; Schunk, 1995; Wigfield & Eccles, 2000; Winne, 1995).

3. Cognitive psychology and psychological-and sociological-constructivist principles of learning, including the implications of these for pedagogy (Anderson, 2000; Bransford et al., 1999; Phillips, 2000).

The importance of understandings

Eisner's (2000) first of twelve key lessons for educators is, "Students learn both less and more than what they have been taught" (p. 343):

> students must … create meaning for themselves and … (these meanings) … are not simply a function of what teachers intend them to learn. Students make meaning from the tools they bring with them … the backlogs of their experiences and the "languages" they know how to use …. (Eisner, 2000 p. 344)

Eisner took a positive view: that the diversity that comes from many students' perspectives is far preferable to a student army all marching to the same drum. Cobb (1994) also proposed that

> there are significant qualitative differences in the understandings that students develop in instructional situations, and that these understandings are frequently very different from those that the teacher intends. (p. 13)

Shunk (1995) stressed the importance of learners' perceptions as mediator effects in the self-regulatory aspects of the learning process:

> What is also important … [for self-regulation] … is the role of learners' perceptions of themselves (e.g. their competencies, interests, values), of others (teachers, parents, peers), and of learning environments (classrooms, environments, homes). These perceptions involve knowledge but are subjective and may conflict with knowledge possessed by learners or others. (p. 214)

Eisner's "tools," Cobb's "understandings" and Shunk's "perceptions" all suggest that learners come to learning situations pre-equipped with thoughts that mediate their learning experiences. I have adopted the broad term *understandings* to describe these thoughts and to focus my further investigations into this field. Of course, it is not only learners who bring understandings to teaching—learning settings: teachers do also (Kerr, 1981; Trigwell, Prosser, & Taylor, 1994).

Prior investigations

In some areas considerable investigations have been undertaken about teachers' and learners' understandings. For example, Perry (1970), Saljo (1979), and more recently, Marshall, Summers and Woolnough (1999), (see also Eklund-Myrskog, 1998; Klein, 1996; Marton, Dall'Alba, & Beaty, 1993; Taylor, 1996) have identified hierarchies of learners' *conceptions* of "What is learning?" Conceptions have ranged from increasing knowledge to change as a person. Another field of investigation has been concerned with people's knowing about knowing, or *epistemology* (Hofer, 2000; Hofer & Pintrich, 1997; Jakubowski & Tobin, 1997; Lampert, 1990; Schommer, 1990; 1993; Schommer, Crouse, & Rhodes, 1992). Hofer (2000) categorised people's epistemological beliefs into four dimensions: certainty, simplicity, source and justification.

A third field of investigation has been *approaches to studying*, perhaps best represented by the work of Biggs (Biggs, Kember, & Leung, 2001; Biggs, 1979; Biggs, 1987) in Australia and work by Entwistle, Hanley and Hounsell (1979) in England. Biggs proposed an interaction between motives -- extrinsic, intrinsic and achievement -- and deep- versus surface- study strategies. Together, motives and strategies become an approach to learning: surface or deep. Trigwell and Prosser (1996) (see also Prosser & Trigwell, 1997; Samuelowicz & Bain, 1992; Trigwell & Prosser, 1997; Trigwell et al., 1994) have conducted similar investigations with teachers' approaches to teaching. For example, a teacher who conceives of teaching as requiring an information transfer/teacher focussed approach may elicit surface learning responses in his or her students. However, a higher level "student centred" theory of teaching

requires teaching methods that have students actively involved in undertaking learning activities and constructing knowledge.

Of course, the perspectives of learners and teachers may differ. In a study of science teaching and learning Tasker (1992 p. 28) drew attention to a *gap* between teachers and students: "Often what I observed was two lessons, the teacher's and the learner's." This suggested possible mismatches between teachers' and learners' views of what a lesson is all about, including its purpose, procedures and outcomes (see also Osborne & Freyberg, 1980; Tasker, 1981; Tasker & Freyberg, 1985; Tasker & Osborne, 1983). Stigler and Hiebert's (1999) analysis of videotapes of mathematics teaching in different nations suggested that not only are there differences between individual teachers and learners, but also that noticeable differences are evident between different cultures' practices and expectations of teaching and learning.

Networks of understandings

Research into conceptions, epistemologies, approaches and teacher—learner congruence has been extensive. However, it has been confined to a fairly narrow range of understandings that teachers and learners might potentially bring to the teaching and learning situation. There are currently many active fields of research in educational psychology and yet there appears to be surprisingly little evidence documented about teachers' and learners' understandings about issues that arise from other contemporary theories of teaching and learning. For example, what do teachers and learners understand about

1. Achievement goals (Pintrich, 2000)
2. Theories of intelligence (Dweck, 1999)
3. Self-regulated learning (Zimmerman, 1998)
4. Metacognition and metacognitive strategies (Flavell, 1979; Nelson, 1996)
5. Help seeking (Karabenick, 1998)
6. Participation in communities of learners (Brown & Campione, 1996; Wenger, 1998)
7. Assessment (Biggs, 1999a; Biggs, 1999b)
8. Motivation (Wigfield & Eccles, 2000)?

Such issues are relevant because they comprise key elements of contemporary discussions about the psychology of instruction and consequent recommendations for educational practice. Furthermore, what do teachers understand about learners' understandings and what do learners understand about the interaction between teaching and learning? Does earlier research, such as Tasker's (1985) gap, still hold true given two decades of explicit teaching pedagogy (for example, see Westwood, 1997), or does a Piagetian perspective (Klein, 1996) dominate? There appears to be scope to extend earlier work by going beyond placing people's understandings along reproductive--constructive hierarchies, or surface—deep dimensions, to investigate more deeply how understandings in areas such as knowledge of strategies for learning *interact* with conceptions of learning. Indeed, it might be that understandings are better conceptualised as *networks of understandings* and my study seeks evidence for this possibility.

Research Questions

If the educational community is to be fully informed about the factors that influence the success of educational programs, then they must have information about the understandings that teachers and learners bring to educational settings. Therefore my research aims to investigate teachers' and learners' understandings about teaching and learning more deeply and widely than is currently represented in reports in the relevant literature. Simply stated, my broad research question is,

"What are teachers' and learners' understandings about teaching and learning?"

Subsidiary questions include

1. "Are teachers' and learners' understandings congruent with each other?"
2. "Are teachers' and learners' understandings congruent with current theoretical perspectives as represented in the teaching and learning literature?"
3. "How are people's understandings best represented—as dichotomies, dimensions, hierarchies, or networks?"

METHOD

Anonymity

All personal and site names used in this paper are pseudonyms.

Participants and Sites

I conducted taped interviews with a purposive (Miles & Huberman, 1994) sample of 10 teachers and 76 learners. These interviews lasted from about 20 to 90 minutes, with the average being about 45 minutes. A few of the younger school students were interviewed in pairs to facilitate their feeling comfortable with the interview situation. I also collected approximately 100 short written responses and some roving, five minute, 'in-class' interviews, so as to access the perspectives of whole classes of school students in single lesson periods.

Figure 1 displays the seven sites, categorised by type of learning institution, and nine classes included in this study. I aimed to access as wide a variety of teachers and learners as possible, while keeping in mind limitations such as time and cost. Thus I approached sites that would give me a range of age groups (grade 3 to adult), levels of education (primary, secondary, further education, pre-entry and graduate entry university) and inner- and outer-metropolitan and rural South Australia.

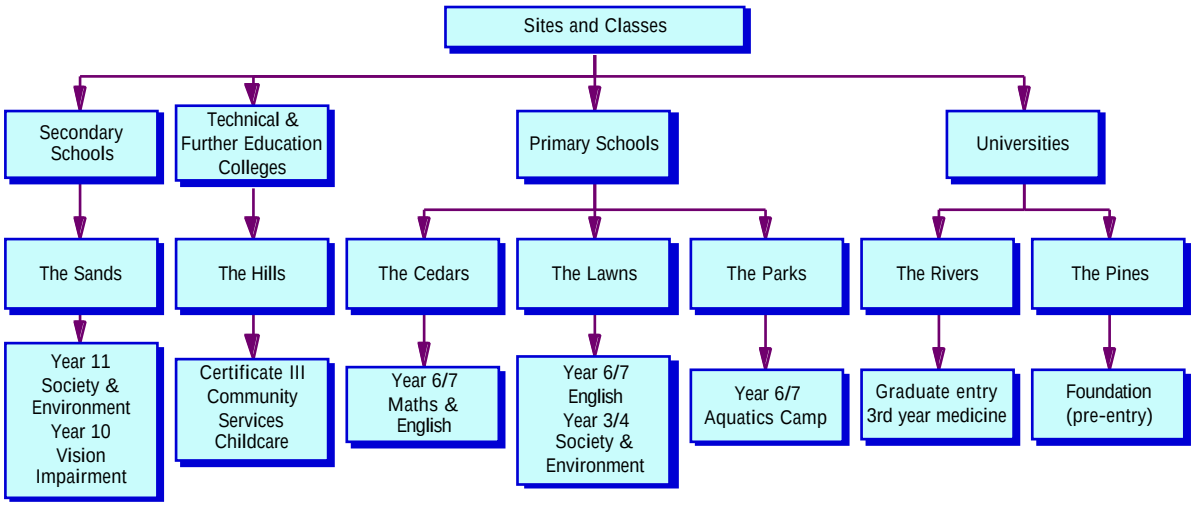


Figure 1: Sites and Classes

Voluntary Participation

After receiving permission from management at each site, I addressed each class group and their teacher(s) verbally and in writing. I outlined the nature of my research and distributed a list of my interview questions. I then requested volunteers to answer my interview questions. All participants gave voluntary consent. I also obtained parental consent for school age students.

Limitations of this study

Voluntary participation raises the question of whether the views of people disinclined to volunteer are different to people inclined to volunteer. Ethically it is not supportable, and practically it is not fruitful, to interview people who do not give voluntary consent. Therefore potential bias in the nature of volunteer responses must be recognised as a limitation of this study.

Another source of potential bias is the interpretive framework of the researcher. Sturman (1999) called for transparency at all stages of data collection and analysis so as to establish credibility of method and results. My data collection, analysis and synthesis procedures follow Sturman's advice.

Further, it is clear that this is a small study of less than 100 participants. The aim is not to make claims that what is interpreted from the data in this study will necessarily apply to other people. Rather, issues of generalisation and applicability are better understood as being a contribution to analysis and theory, as explained by Miles and Huberman (1994, p. 27-28) in their study of role modeling in a kindergarten:

> Sampling like this, both within and across cases, puts flesh on the bones of general constructs and their relationships. We can see generic processes: our generalizations are not to "all kindergartens," but to existing or new theories of how role modelling works. As Firestone (1993) suggests, *the most useful generalizations from qualitative studies are analytic, not "sample to population."* (my italics)

This issue was also considered by Cobb (2001 p. 549-460) in his discussion of the generalizability of design experiments: "This is generalization by means of an explanatory framework rather than by means of a representative sample."

Interviews

My aim was to comprehensively capture each participant's understandings about their own teaching and learning *in action*. This was based upon my recognition of the importance of all three vertices of Bandura's (1997) cognition, environment and behaviour triangle of social-cognitive theory, as well as heeding the role that context plays in teaching and learning (Lave, 1988; Wenger, 1998). Therefore the explicit focus of each interview was each participant's current lesson, topic and/or course. I used the interview questions listed in Table 1 to guide the direction of each interview. I also added extra probing questions according to the idiosyncratic direction that each interview took.

The interview data was supplemented by my non-participant observation and note-taking of each teaching -- learning setting so as to provide contextual information that would provide a background to the interview data.

Interview Questions.

Founded in the theory introduced on page 294, I composed a set of 18 interview questions. The interview questions are not exhaustive, given practical limitations, but are extensive and seemed manageable. Table 1 displays those questions. Column 1 indicates the theoretical field that precipitated each question. Column 2 lists questions designed to elicit information from teachers about their understandings in each field. Column 3 lists matching questions for learners.

Table 1: Interview questions

BACKGROUND THEORY	QUESTIONS FOR TEACHERS	QUESTIONS FOR LEARNERS
Goals	What do you want your student/s to achieve from what they are doing in this lesson/topic? Why?	What do you want to achieve from what you are doing in this lesson/topic? Why?
Theory of intelligence	How well do you expect your student/s to perform in this lesson/topic? Why? Can that be changed?	How well do you expect to perform in this lesson/topic? Why? Can that be changed?
Metacognition	What thinking processes will students be using in this lesson/topic?	What thinking processes will you be using in this lesson/topic?
Self -regulation	In what ways are you responsible for the learning in this lesson/topic? In what ways are your students responsible?	In what ways are you responsible for the learning in this lesson/topic? In what ways is your teacher responsible?
Assessment	How will you know that your student/s have learned what they are meant to?	How will you know that you have learned what you are meant to?
Objectives	What specific things do you want your students to learn from this lesson/topic? What broad understandings or ideas do you want your student/s to get from this lesson/topic?	What specific things are you meant to learn from this lesson/topic? What broad understandings or ideas do you think you are meant to get from this lesson/topic?
Purpose	Why are you teaching this? When, where and how will your student/s use the learning in this lesson/topic?	Why are you learning this? When, where and how will you use the learning in this lesson/topic?
Strategies	How will your teaching and learning strategies help your students to learn?	How does what you are doing help you to learn what you are meant to?
Value and Interest	Is this what your students want to learn? Why or why not?	Is this what you want to learn? Why or why not?
Communities of learners. Strategic help seeking.	Who and/or what helps your students to learn? How?	Who and/or what helps you to learn? How?

DATA ANALYSIS

I have not yet completed my analysis of all of the interviews, therefore the data reported herein is provisional and part of my work in progress. Interviews were transcribed verbatim. I employed NUD*IST 4 (QSR, 1997) data analysis software to categorise and code the interview transcripts. I then exported the NUD*IST category trees to Inspiration© (Helfgott & Westhaver, 2000) display software.

Categorisation of Units of Meaning

I categorised each *Unit of Meaning* in each transcript using a combination of *deductive* categorisation and *inductive* category formation (Miles & Huberman, 1994). The deductive categories are grounded in the three foundations that underpin this study as outlined on page 1[1]. I labeled the deductive categories "Domains" (5) and "Constructs" (10). These are represented in *Figure 2.*

[1] In brief, 1) a philosophy of intentions, plans and actions; 2) social-cognitive theory including motivation and self-regulation 3) cognitive psychology and constructivism.

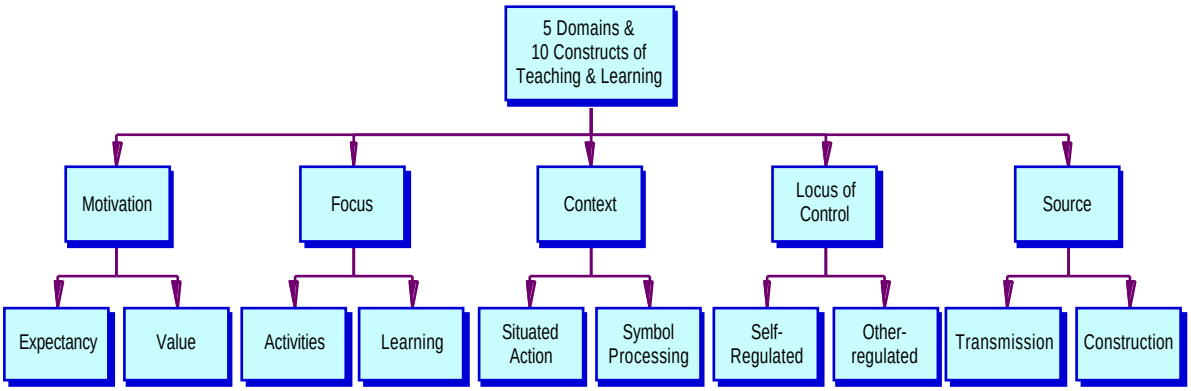


Figure 2: Domains and Constructs in Teaching and Learning

Figure 2 includes, from left to right; expectancy—value constructs in motivation (Wigfield & Eccles, 2000); a focus upon activities versus a focus upon meaningful learning (Bereiter & Scardamalia, 1989); the contexts provided by situated action and symbol processing (Anderson, 2000; Wenger, 1998); self- and other-regulated loci of control (Zimmerman, 1998); and constructivist versus transmissionist philosophy and pedagogy (Bransford et al., 1999; von Glasersfeld, 1995). Together these broad Domains and Constructs provided a comprehensive coverage of the Units of Meaning in the hundreds of pages of transcripts obtained in my study.

Figure 3 illustrates an example of the categorisation process for one small section of Rory's transcript. Reading from left to right, I first divided Rory's transcript into Units of Meaning. One of those units of meaning was "Learn by being here" which I will use to illustrate each step of the categorisation process. It can be seen that I allocated "Learn by being here" to the Domain: *Context* and the Construct: *Situated Action* (from *Figure 2* above). I then created the inductive Theme "Learn by being here" which reflects Rory's own words, and next I included Rory's Elaborations of that Theme, such as "osmosis." I then added the Elaborations "watching" and "asking questions." Had Rory elaborated further, I could have continued to expand the figure on the right hand side.

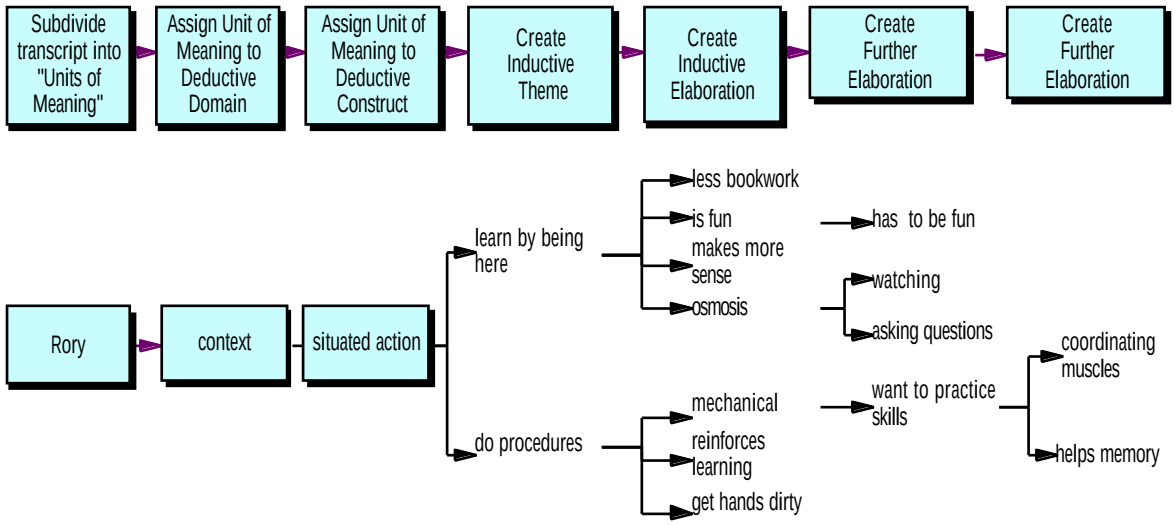


Figure 3: Example of categorisation

The detailed categorisation and display of data discussed above worked so as to, in a qualitative sense, "standardise" the data. Clearly this standardisation is not in the quantitative sense of, say, calculating Z scores. However, the uniform processes of categorisation and display brought the understandings of participants into a common framework, enabling accessibility and interpretation. This procedure also makes my data analysis procedures explicit, thus allowing readers to audit the veracity of any interpretive claims I might make.

RESULTS AND DISCUSSION

Figure 3 displays the process I undertook for analysing and displaying each participant's transcript. The process resulted in 10 or more displays for each participant, thus enabling comparisons and contrasts. I will turn now to a synthesis of the data obtained from my analysis of the transcripts.

Integration of the Ten Constructs of Teaching and Learning

My major interpretation of participants' understandings is that participants appear to *integrate* the ten Constructs of Teaching and Learning that I have used as theoretical foundations for this study. For example, Sally (medical student) indicated that she was confident in her own ability, that she had a busy brain that made fast connections, and that she wanted to impress other people (entity theory/performance goals). She also told how she invested a lot of effort into her study and could become frustrated with herself (incremental attributions/mastery goals) (Dweck, 1999; Dweck, 1986; Pintrich, 2000). Sally takes responsibility for her own learning while recognising the role that other people play in informing her about what is expected of her (Bransford et al., 1999). Sally knows that there are various sources of knowledge that can be transmitted to her, but she is not passive in this process. She has a range of strategies that facilitate her taking an active role in knowledge construction, including "having a go" and knowing that in doing so it is "OK to be wrong." She uses surface and deep strategies for learning (Biggs et al., 2001) and holds low and high level conceptions of learning (Marshall et al., 1999). Sally holds extrinsic (future employment prospects) and intrinsic (interest and personal fulfilment) motivations (Murphy & Alexander, 2000).

This integration of constructs also seems apparent in the transcripts of other participants. To interrogate this integration further, I created a Construct X Construct[2] matrix from the medical cohort's transcripts (7 students, 1 mentor) which is reproduced as Table 2. Each cell of the matrix highlights a Unit of Meaning that suggests integration of two Constructs.

Table 2 illustrates, for example, that the statement "complement experience with book learning, see first, read first, doesn't matter" can be interpreted as an affirmation of the positive benefit of situated action and symbol processing working together to promote learning. Another example, "learning is a shared responsibility between student and supervisors" illustrates an interaction between self- and other-direction. A third example, "information comes to you -- ask questions" illustrates an interaction between transmitted information and constructing knowledge. The presence of numerous interactions in participants' understandings leads me to begin to raise questions about dichotomies that appear in the literature. I will discuss some of these interactions in further detail under separate headings below.

[2] The ten Constructs are displayed in *Figure 2.*

Table 2: Medical cohort -- Construct X Construct interactions

	expectancy	value	activities	learning	situated action	symbol processing	self	other	transmission	construction
expectancy										
value										
activities	cramming before an exam	some things not relevant/interesting								
learning	confident to do alright _have self doubts	be a good doctor	exam is small part of learning medicine							
situated action	not an effort to learn	like going to work - exciting-feel alive-useful		osmosis-learn by being there						
symbol processing				provides foundation knowledge	complement experience with book learning-read first-see first-doesn't matter					
self-direction	I can do this			integrate learning experiences with curriculum guide	put myself in the position that I'm getting the experience I want	sit down and read				
other-direction	please my GPs		supervisors think some things are important-I don't	inspiring people	mentor checks your work		learning is shared responsibility between student and supervisors			
transmission				lectures-the facts	learning from expert	books present information nicely	.listen-watch-be there	tutors impart knowledge from their experience-terrific to learn that from them		
construction		assimilate a broad understanding treat patients better		join categories together-new understandings open up	patient presentation raises a question in your mind	draw pictures & diagrams-pattern learn	try to make sense of it-what does it mean	be questioned-be challenged	information comes to you-ask questions	

Approaches and Conceptions of Learning

Biggs (Biggs et al., 2001; Biggs, 1979; Biggs, 1987) argued that students will adopt deep or surface *approaches to studying* in interaction with the affordances of the learning environment. Participants in the present study certainly appear to hold deep and surface approaches to studying. However, whereas Biggs et al. (Biggs et al., 2001 p. 142) proposed that deep and surface approaches are, "negatively correlated since deep and surface approaches are envisaged as not commonly occurring in conjunction," the participants in this study appear to be able to hold deep and surface approaches *in conjunction* and to utilise both approaches as valuable resources for learning. These "in conjunction" approaches seem to apply to both the strategy and motive sub-components of deep and surface approaches. This is in contrast to some of Biggs et al's findings: for example, that surface strategy and deep motive covary negatively (2001 p. 144).

Memorising. Figure 4 displays how Johanna (medical) told how she "swats like mad" as part of accumulating knowledge, and that swatting provides her with the information she needs to allow "new understandings to open up." Swatting might be important to combat forgetting, because "some goes away," (Nuthall, 2000a; 2000b), but Johanna understands that when the time comes for her to "use it again" she will "remember more" and that she will not be "trying to understand it the next time." Nor does she have to remember it all, as long as she understands the principle. Most interesting is Johanna's comment that the plan is to "have it all in your head at once," that is, the swatted knowledge, the broad understandings, and the real life. My interpretation of Johanna's complete transcript is that she has deep motives for learning. The Units of Meaning in *Figure 4* suggest that Johanna is adopting deep strategies for learning, and also, that one of her key strategies--swatting--is surface.

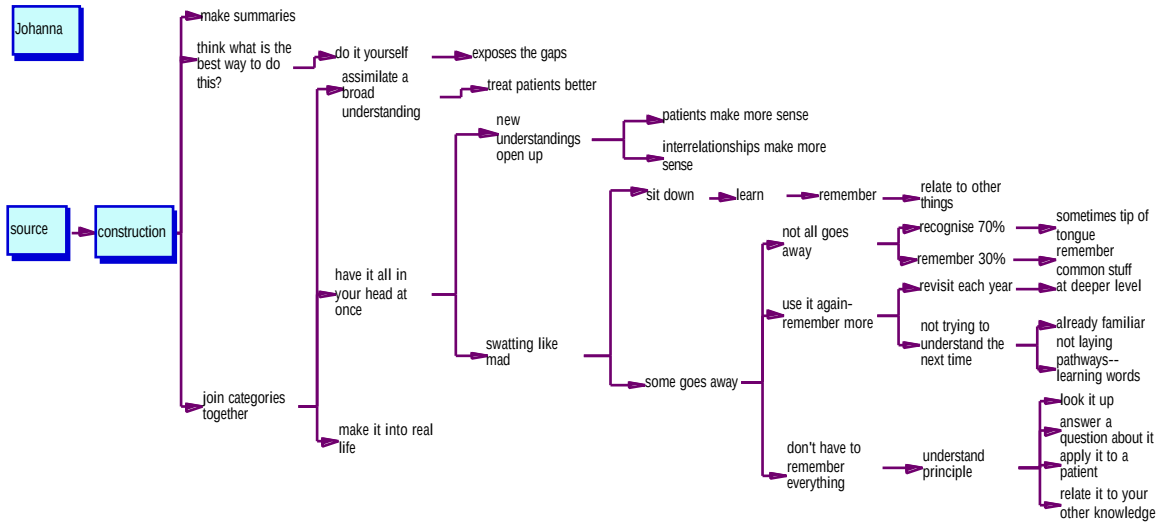


Figure 4: Johanna--construction

Participants spoke of other strategies that could be categorised as surface, such as reading over and over, repeating without looking, memorising and practising the items on competency check lists. As in Johanna's case, such surface strategies were integrated with deep strategies such as drawing flow charts and having discussions as, well as with deep motives, such as "be a good doctor" or "be a good child care worker." It is interesting to note a remarkable similarity between Sally, in her 6th year of high achieving university studies, and Amber, in her 6th week of Foundation level university pre-entry studies: Sally had pasted to her kitchen wall, and Amber had pasted to her bedroom ceiling, pictures of what each was trying to memorise. School students did not seem to be as concerned with understanding and remembering information, but instead spoke of reproducing and presenting information.

Base Knowledge. In part, the need to memorise appears to stem from participants' understandings about the importance of acquiring extensive quantities of base knowledge necessary for further development in their field. Indeed, John lamented that he was a "hopeless rote learner." He saw rote learning as a valuable resource for acquiring essential information necessary to be a good doctor -- a deep level motive. Many participants spoke of the need to accumulate a 'base' or 'foundation' of knowledge. This knowledge is not the kind that they could work out for themselves, but is what needs to be learned from authoritative sources. It is also the knowledge needed to enable deep thinking about issues such as case management, dealing with children, or constructing an academic argument.

Memorising has been identified in the literature as a lower-level learning conception and as a strategy associated with surface motives for learning. Marshall et al. (1999) placed memorising at level A (lowest) and Marton et al. (1993) placed memorising at level B (second lowest) in their hierarchies of conceptions of learning. Furthermore, Biggs et al's (2001 p. 148) Study Processes Questionnaire items that refer to memorising and rote learning are surface strategy items:

> 8. I learn some things by rote, going over and over them until I know them by heart even if I do not understand them
>
> 11. I find I can get by in most assessments by memorising key sections rather than trying to understand them.

An alternative view might be that learning by memorising is a valuable cognitive strategy for dealing with extensive quantities of base knowledge, with the eventual aim that such base knowledge will provide the necessary building blocks for deeper-level learning.

Teaching by telling (transmission). To take my synthesis of participants' understandings further, participants explained how information that was "told" to them by other people (texts,

lecturers) was considered to be equally as valuable as information they constructed from their own experiences.

> I think they have a responsibility to impart the stuff to us that isn't in the textbooks, the sort of practical stuff they've gleaned over the years, like how to examine a screaming baby. You know how to examine the baby but if it's all upset, you can't hear its heart and all that kind of thing, what do you do? That sort of practical stuff is just really terrific to learn from them ….

Teaching by telling could be seen as a transmissionist methodology that evokes a surface approach in learners (Trigwell, Prosser, & Waterhouse, 1999), as opposed to active involvement in hands on, minds on learning activities (Tobin, Tippins, & Gallard, 1994). However there are many instances in the transcripts where participants who could be considered to be adopting a deep approach to learning, or operating at a high level of conceptions of learning, understand teaching by telling to be an essential resource for deep learning. For example, *Box 1* tells what Johanna had to say about the responsibility of tutors.

Box 1: Johanna--teaching by telling

Learning by "absorption." Conceptions of learning theorists, such as Marshall et al. (1999) and Marton et al. (1993) have proposed a qualitative distinction between lower- and higher-level conceptions. For example, Marshall et al. and Marton et al. rate "absorption" as a low-level conception. However, many medical and child care students value absorption, or in alternative terminology, "osmosis" (in the context of situated action) as a way of achieving deep-level learning goals. An alternative interpretation to that proposed by Marton et al. and Marshall et al. might be that absorption could be classified as a high level approach, for example in gaining practical intelligence (Sternberg, 2000), in gaining tacit knowledge (Edwards & Mercer, 1993), in learning about the contexts of action (Wenger, 1998), and in the ways that different cultures teach and learn (Stigler & Hiebert, 1999).

Approaches and conceptions of learning summarised. Previous 'approaches to studying' interpretations have tended to indicate that the person 'adopts' or is 'located at' a particular place on a learning dichotomy in interaction with a particular teaching and/or assessment style, and, furthermore, that surface approaches are inferior to deep approaches. For example, Biggs et al. (2001 p136) stated

> A student, who typically picks out likely items for assessment and rote learns them, finds that strategy won't work under portfolio assessment, so goes deep. Another student who normally interacts deeply, may decide to go surface in a module that is overloaded with content and assessed by a poorly constructed MCQ. Indeed the generic aim of good teaching is precisely to encourage students to adopt a deep approach and to discourage the use of a surface approach (Biggs, 1999).

My interpretation of participants' understandings suggests that it may be fruitful to seek a more complex representation of surface and deep approaches to learning. One possibility is that learners can exploit surface level strategies whilst motivated by the pursuit of deep level goals, or higher level conceptions of learning. A second possibility is that the decision to adopt surface or deep approaches to learning is influenced by an interaction of factors over and above the method of teaching or the format of assessments. For example, adult participants spoke about the interaction between the volumes of information to be mastered in the limited time available. Even Amelia, in Year 4, when asked, "What was happening in your head [in the lesson]?" replied, "Hurry, hurry, hurry." Time is a realistic constraint. Also, many participants referred to motivations such as interest, employment and previous life experiences as influencing what they were prepared to do to learn. This places approaches to learning into a much broader framework than the affordances of the existing teaching -- learning situation.

The issue seems to be the *level* of analyses. It is uncontroversial to say that at the level of a whole course of study, a deep level approach, and learning leading to a changed conception of oneself in interaction with the world, is desirable. However, some of the participants in this study seem to suggest that at a task level, surface level strategies may not compromise deep level motives. This interaction between surface strategies and deep motives might provide an explanation for the variable results that Biggs et al. (2001) found in isolating an "achievement" approach to learning. Furthermore, in practical terms of class size and teaching resources, there may be some benefit in assessing exactly which learning can be profitably approached using surface strategies, so as to free up resources for deep learning in other domains. It seems that some participants in this study are indeed making that assessment for themselves.

As well as approaches to studying, there seems to be scope for re-assessing existing descriptors of levels of conceptions of learning. Memorising and absorption are considered by many participants in this study to be valuable strategies to enable goals of "being a good doctor/child care worker." The value of such strategies for achieving goals of dealing with substantial quantities of information, and in learning about the intricacies of complex situations, could be reassessed.

Goals

As with my preceding discussion about approaches to learning, my analysis of participants' interviews suggests that it may be facile to speak of mastery (deeply learn the material) and performance (show other how much you know) goals as a dichotomy (Dweck, 1999; Pintrich, 2000), as such goals appear to be able to operate simultaneously. *Figure 5* displays Sally's account of how her mastery goals of 'learning medicine' and 'being a good doctor' incorporated her performance goals of passing exams.

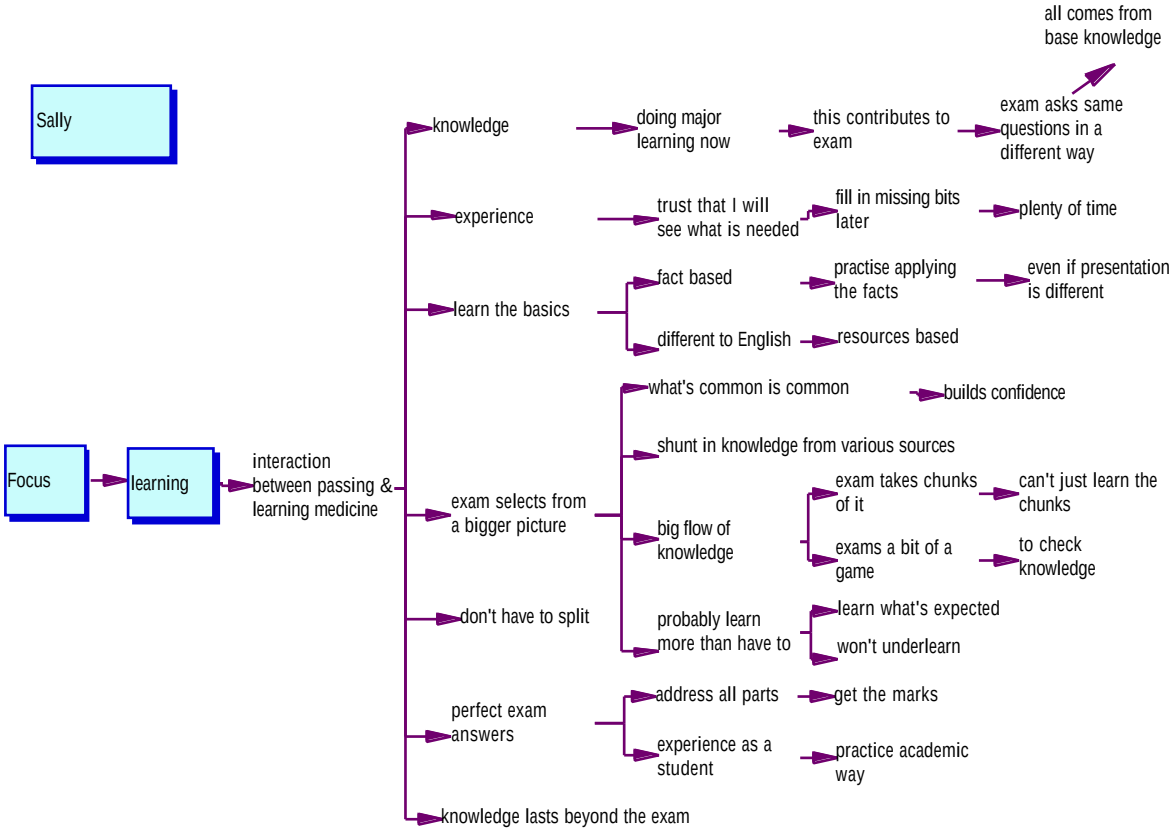


Figure 5: Sally--passing exams and learning medicine

Sally referred to the 'perfect exam answers,' 'address all the parts' and 'get the marks'. She also told of how the exam 'selects from a bigger picture' and that the knowledge 'lasts beyond the exam.' She said you 'don't have to split' exam knowledge and learning medicine - it is all the same knowledge. It is interesting to compare *Figure 5* with *Figure 6: Roxanne--focus—learning.* Roxanne also integrates passing exams and being a good doctor, however she adds an important caveat: that the two should go hand in hand, but don't due to limited time.

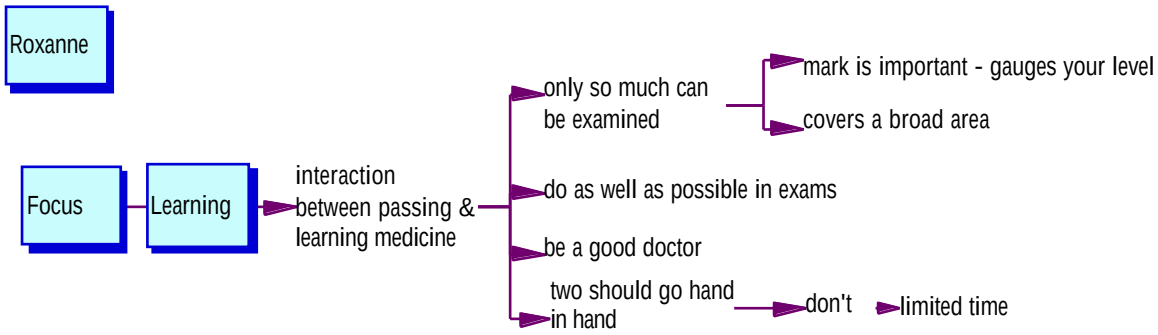


Figure 6: Roxanne--focus—learning

The child care students also described what they do to pass assignments (read and copy) compared to what they do to care for children – either in the childcare centre or with their family and friends ("think about what to do to meet the children's needs"). A possible application of achievement goal theory to the medical and child care students is that performance and mastery goals seem to work together in a complementary fashion rather than separately or as alternative perspectives. This seems to bear similarity to Zimmerman and Kitsantis' (1997) morphogenic goal theory, where mastery and performance goals come into play at different stages of the learning trajectory. Alternatively, Pintrich's (2000) achieving goal seems to capture the flavour of participants' mixing of the two constructs. A few students also differentiated performance goals into two: showing other people what you can do and showing yourself what you can do by way of passing external criteria such as exams.

Interviewees also highlighted the practical implications of passing or failing exams, thus adding many more parameters to decisions about adopting mastery or performance goals than might prevail in school classrooms or laboratory simulations. For example, Antoine (medical) warned of the dangers of learning certain things, such as one's special interests, too deeply. Why? Because such an approach would steal time from other essential learning. In the end, no matter how deeply a student learns about his or her special interests, s/he can't be any sort of doctor, let alone a good doctor, without passing the medical exams.

Amelia: Year 4
There was this thing in the hall -- a debate -- and it was on, 'we need homework or not.' And Madeline said, 'why do we need it -- it's easy -- I learnt this when I was little: cut -- copy -- paste -- print. And you get it off{the computer} and just stick it straight in your book and you don't read it. So if you copy it out yourself and out in your own words -- you understand -- it's in childrens' words, and it's sort of a bit easier for Mr H -- maybe not to read, but say we get a bit more points usually because you've read through it, and put it in your own words and you've done more work.

Box 2: Cut-copy-paste-print

A different picture emerges from the school students' transcripts. Many school students conveyed a sense of a performance orientation when they spoke about school work, but a mastery orientation when they discussed their out of school, personal interest activities. *Box 2* gives Amelia's description of a performance orientation to project work. Freda (Year11), Antony (Year 6), Amelia (Year 4) and Hanna (Year 4) told of how project work meant gathering together and re-presenting information.

Interviewer: So you don't play around or try different things?
Donna: You usually get sort of set things that you have to do - so can't really sort of change it too much cos you've got things that you have to do.

Box 3: Set things

However, when discussing her flute playing, Freda used words such as "invent," "fun" and "make new tones." Antony spoke of his drawing (at home) using words such as "kept trying and trying." Donna said that in learning how to use make-up, "You just sort of play around and see what works" and "Try different things." But that is not what happens in Australian Studies according to Donna's account in *Box 3*. These examples can be related to the work of Bereiter and Scardamalia (2000; 1989) who lamented that schoolwork often focussed upon activities and products (such as projects) at the expense of deep learning of subject matter.

It is enlightening to observe that school students do have understandings about strategies that help them to learn. They also appear to have understandings about when they will, or will not, adopt deep or surface strategies to achieve mastery or performance goals. This supports researchers' and practitioners' calls for creating institutional learning environments that maximise the potential for optimum teaching and learning behaviours (Biggs, 1999b; Brown & Campione, 1996; Paris & Ayres, 1994).

Theories of Intelligence

Dweck (1999) discussed people's disposition to adopt a malleable (I can get better at this) or fixed (this is my level) theory about their intelligence. The medical student Rory said that he was not super-intelligent; that he didn't work really hard; that he procrastinated; that he expected to do well in exams and had always done so; that he enjoyed exams as it was a chance to demonstrate to himself and others what he knew, and; that he was anxious about exams because he wanted to do well. This part of Rory's account identifies him as a fixed intelligence theorist. However, Rory also said that, as there was no need for him to perform well in exams (as his intern placement was already established) he was free to attend to his main medical interests; to have fun with his learning; to know his own deficiencies; to work out what he needed and wanted to learn; to seek out opportunities for learning; to be willing and available to learn; to establish good relationships with General Practitioners and staff to enable learning; to chase up learning; to monitor his own success using a framework of learning goals, and; to make notes, read up and think through what has happened. These deep learning strategies align with Dweck's description of what a person with a malleable theory of intelligence would do.

Rory was not alone. Other participants spoke about their learning as if they held both fixed and malleable theories. The interaction between the two theories is well illustrated by the Unit of Meaning "work at it to the best of my (their) ability" which was spoken by many teachers and learners. "Work at it" flows from a malleable theory of intelligence. "To the best of my [their] ability" flows from a fixed theory of intelligence. The complete Unit of Meaning implies and demands both. Note that this statement also hints at another kind of ability – the ability to apply effort, which was suggested by Stigler and Hiebert (1999) in their discussions about Japanese students' attributions for success or failure.

The most telling evidence for the variable influence of other factors upon theories of intelligence was provided by the Foundation students. Anne (Foundation) told of her previous understandings about her own ability and application of effort, summed up by her father's oft repeated taunt, "You're a bugger of a child." Anne said she wasn't a bugger of a child any more: she had the ability and could apply the effort. Amber (Foundation), who experienced similar demeaning feedback from her father as a child and as an adult, indicated that she was discovering that she appeared to have previously unrecognised abilities and, supported by her

husband, was now willing to apply as much effort as necessary to make the most of her ability. Ray (Foundation) told of how school never interested him, how he did poorly and how he left at the first opportunity. Now he has decided that life has more to offer than a succession of itinerant, physically demanding jobs. Although English is his second language, and although he has never written a page let alone a whole essay, Ray's understanding was that he wanted to learn, that he presumably had as much ability as anyone else, and he was prepared to "do what it takes" because he "wanted to"[3]. My interpretation is that participants hold much more complex understandings about their intelligence than is captured by the duality of "fixed" and "malleable" theories. From the transcripts, other interacting factors include self-efficacy in relation to the specific task or domain, the availability of a key support person, interest, desire, necessity, the freedom to choose to learn and the affordances of the teaching--learning environment.

As with approaches to- and conceptions of- learning discussed above, fixed and malleable intelligence theory seems to illuminate parts of, but not the complexity of, people's understandings. The danger, although it may never have been the intention of the authors, is that such theories may be interpreted by researchers and practitioners in a static, dualistic and individualistic fashion. For example, Biggs et al. (2001) stressed that the Study Processes Questionnaire (SPQ) is best used to measure, and even modified to fit, specific teaching contexts. Nevertheless, I suspect that the tendency would be to interpret responses to the SPQ to indicate that *individual respondents* occupy a surface or deep position which, although it may change over the duration of a course of study, does not change from task to task or situation to situation. Certainly the wording of some SPQ items encourages an individual rather than a contextual focus. Similarly, Dweck (1999), although proposing that individuals hold a fixed or malleable theory, demonstrated how it was possible to experimentally manipulate people's theories of intelligence with as little as one intervention, thus clearly accounting for the influence of context upon people's fixed or malleable view. My assessment from the interviews in the present study is that the explanatory frameworks of dualistic theories pay insufficient attention to the interactions that occur within each individual's understandings and between each individual and each situation.

Situated Action

Many participants described the difference in effort required when learning in a situated or, in more common terminology, 'hands-on' learning environment, compared to learning from books or lectures. Possible reasons for learning in situated action making learning less effortful include that

1. Additional contextual cues are provided by sights, sounds, touch, emotions and social interactions.
2. Contexts promote and provide opportunities to construct and ask questions.
3. Students may lack essential metacognitive skills for text-based learning.
4. Hands-on activities might work towards precluding a 'surface approach' to learning. For example, it may be less possible to 'skim over' real time experiences as it is possible to 'skim' a text. Also, it may be less possible to disengage from the sensory and emotional contexts of human interaction than it is to disengage from text. Thus a deep, rather than surface, approach to learning might be enabled by a situated approach.
5. Opportunities are provided for practice but with slight variations in context with each repetition. Note that such opportunities for practice are not restricted to simplistic interpretations of motor skills such as suturing or nappy changing, but can refer to

[3] Ray was awarded a distinction for his first ever university essay.

complex aspects of situated practice, such as developing client relationships, applying mental models, selecting follow up investigations and proposing solutions.

A potential explanation is that learning in situated action stimulates multiple sensory inputs, which in turn stimulate diverse cognitive activity. This diverse brain activity helps to overcome problems of inaccessibility of inert (Whitehead, 1942) unconnected (Anderson, 2000; Bereiter, 2000; Lakomski, 2000), non-transferable pieces of information (St Julien, 2000) and facilitates remembering.

My assessment is that learning in situated action looms as a major issue in participants' understandings. In an attempt to investigate this further, I constructed Table 3, which amalgamates the medical cohort's understandings about situated action. An analogy can be drawn between Table 3 and a quantitative factor analysis. Each column is dedicated to a participant. Each row identifies common Units of Meaning (variables) that appear in the medical cohort's transcripts. I have allocated an interpretive title to each row (or factor). I have then placed the 11 interpretive titles (factors) under the common umbrella of the higher order factor "Situated Action."

Table 3: Medical cohort -- situated action

	Context-- situated action							
	Rory	Sally	Antoine	John	Roxanne	Troy	Johanna	Dr Be
Mentoring		expert checks you	better access to doctors		get GP perspective	one on one experience	one to one with GP	mentor shows student
Apply to real life	reinforces what you've learnt already	see patients-not in textbook way-apply medical model	apply theory to patient		seeing things in real life	patient gives symptoms-you think and apply knowledge	apply theory to patient	connect base knowledge to patient presentation
Osmosis	osmosis-watching asking questions	osmosis theory-learn by being there	see things-don't realise how much I've learnt		exposed to it	seeing patient better than book learning	absorb-not an effort to learn	easier to learn when you see things
Apprentice-ship	learn by being here-makes more sense	have a go-like an apprentice-ship	broad experience-remember people-relevant	perform in the role-like an apprentice-ship		put me in the environment-allow me to train-like apprentice-ship	like an apprentice-ship--don't have to sit down & learn it	
Affect	fun	on job training-can't imagine any other way	see things-affects me	like it more	learn the whole process - more three dimensional	Like a worker-patient forces you to be active-it's energising	like going to work--feel alive	throw in at deep end
Practise	practise skills	practise					practise	
Have a go	do procedures-get hands dirty-reinforces learning	get hands dirty	hands-on	good with my hands	get hands-on experience	hands-on	get practical experience that isn't in books	do things
Safety net	always under supervision	GP vets work	supportive environment			GP always there to check		let student express his/her opinion first
Take Responsibility		work with the patient myself	you're the assistant	make own decisions-challenged	being responsible for all aspects of being a doctor		Think-what is the best way to do this	examine patients-diagnosis
What's common	what's common is common	what's common is common	what's common	common problems	common & rare things		pick up on common things	what's common also rare things
Unstructured		whatever comes through the door	whatever procedures are happening for the day	learn what comes in the door		you just get whatever walks through he door	whatever comes in the door	depends on what comes through the door

The interpretive titles for each row of Table 3 identify key features of learning in situated action, including; the guidance of a mentor; the opportunity to apply your knowledge to real life situations; the apparent ease of learning in practice; the comparison with an apprenticeship model; the positive affect created by the learning situation; opportunities for practice; having a go; the need for a safety net provided by more experienced staff; the opportunity to take responsibility for making one's own decisions; that everyday practice illuminates what is

common (and therefore necessary to know), and; the unstructured nature of the learning environment.

In particular, many participants indicated that learning in situated action is easier, as is captured by *Figure 7*: Jenny (childcare) and *Figure 8*: Johanna (medical). Jenny doesn't have to "shove the learning in" and for Johanna it is "not an effort to learn."

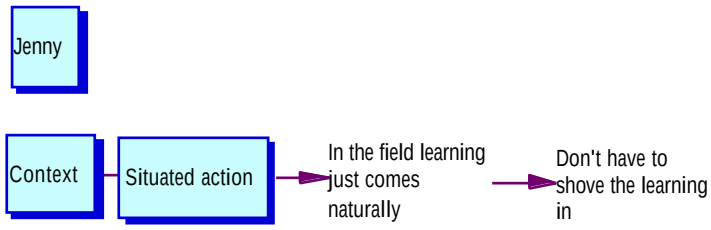


Figure 7: Jenny-situated action

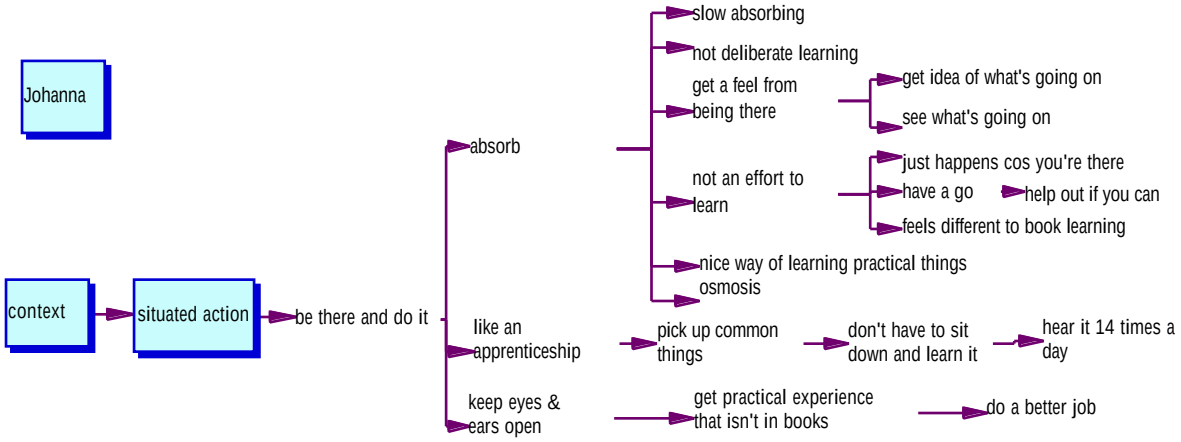


Figure 8: Johanna—situated action

Indeed, Anderson (1987 p. 473) argued that "our learning mechanisms are adapted to deal with complexity. In seeking laboratory simplifications, one may have thrown out the features that would reflect these adaptations to complexity." My hypothesis then, is that situated action embeds knowledge in the complexity of meaning, or pattern, which is another way of saying establishing multiple neural connections, thus making learning less effortful.

Integrating Theory, Practice and Reflection

An interesting aspect of two of the cohorts in this study -- medical and childcare -- is the course(s) design, that juxtaposes opportunities to learn from situated action and opportunities to learn from books (and other texts). Roxanne (medical) described how learning from texts gave her a two-dimensional kind of knowledge. However, when she added her patients to her two-dimensional knowledge she created a much more rich three-dimensional knowledge structure. An impression that emerges from many medical and childcare participants' transcripts is that the temporal combination of theory and practice multiplies to facilitate learning more than a simple addition of the two parts. All of the medical students spoke about the interaction between theory and practice. For example, Roxanne and Troy said that real life situations identify what is important and so directs their reading and John described how he would sit in with surgeons and read up about it later. Childcare students and their teacher told of how they may deal with an anxious child as a learning issue in a lecture session and within a few days be called upon to deal with such a child in the childcare centre.

Figure 9 displays Lauren's (childcare) account of the integration between learning from books and learning in practice. Lauren describes how "putting what they tell you into practice" is "not just sitting getting the information" but "taking it in and doing it," allowing her to "see that you can do it". She says that "books help you with your work" and that the course is able to "give me the information I need and the opportunity to put it into practice." Clearly theory and practice are essentially integrated in Lauren's understanding about effective learning.

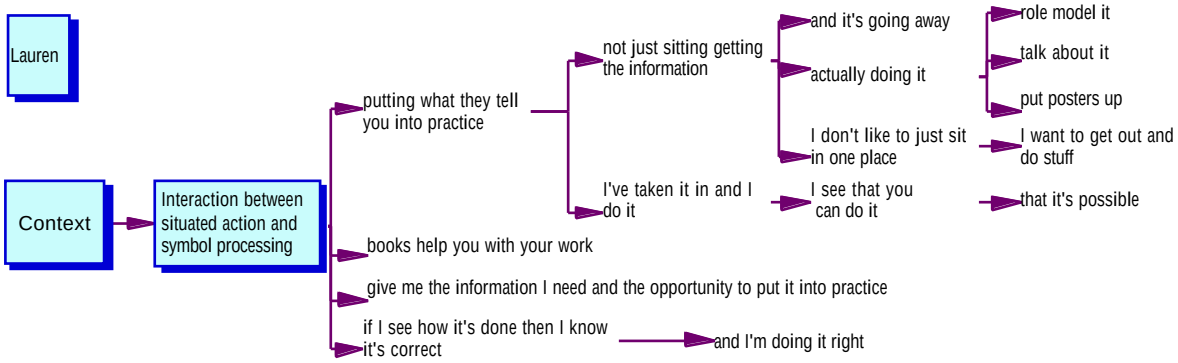


Figure 9: Lauren--Interaction between situated action and symbol processing

Possible reasons as to why the temporal combination of the theory and practice might work so well include that

1. The practical and text experiences provide multiple exposures to the material to be learned: an essential precondition for learning as described by Nuthall (2000a; 2000b)

2. The structure of the learning environment allows a combination of practical experience with time for reflection and follow up of learning issues. Time for reflection, as opposed to another round of frantic practice, is an essential pre-condition for constructing knowledge (Alagumalai, 1999; Baird, Fensham, Gunstone, & White, 1991; Bruning, Schraw, & Ronning, 1999; Day, 1999). A combination of theory, practice and reflection also accords with Hill's (1999) recommendations (albeit Hill referred to teacher education). Reflection is also an essential component of self-regulated learning (Schunk & Ertmer, 1999; Zimmerman, 1998).

Koschman (2001) (after John Dewey) called for a transaction between learning as individual acquisition and learning in social practice. The beneficial integration of individual study of theory, legitimate practice in a like minded community and time for reflection, as described by participants in this study, might provide a model of how such a transaction can be facilitated in teaching -- learning settings.

CONCLUSIONS AND FUTURE DIRECTIONS

This paper has described an "in progress" interview study that seeks answers to the broad research question, "What are teachers' and learners' understandings about teaching and learning?" I have reported only a small fraction of many hundreds of pages of interview transcripts collected for this research. However, this paper begins to give a description of the broad and deep understandings that teachers and learners hold, including understandings about motivation, learning in situated action, self-direction of learning and constructing learning.

My first subsidiary research question was whether a 'gap' appeared between teachers' and learners' understandings. In general, I have not found evidence of a gap between teachers' and learners' understandings about the purposes of their joint endeavours. Although there may be

task-specific, short-lived misunderstandings, or disagreements between teacher and student as to what is relevant to learn (such as John's annoyance with the professional development module of his training), the overall purposes of learning (to get the knowledge required to be a good doctor/to be a good child-care worker/to learn how to prepare an academic argument/to prepare for adulthood) were shared between each teacher and their students. Even questionable or vague purposes (We're doing this in primary school so we can do it in high school) were explicitly shared.

The second subsidiary research question was whether there is congruence between teachers' and learners' understandings and contemporary literature in educational psychology. Whereas participants identified many issues that are currently represented in the literature, my assessment is that participants appear to hold understandings in interaction, especially in interaction with context. Examples include the exploitation of surface strategies for deep learning; the interaction between fixed and malleable theories of intelligence; the potential to concurrently hold mastery and performance goals; the relative ease of learning in situated action, and; the apparent multiplicative learning effect of integrating individual learning from texts with learning in the social contexts of situated action. Participants have provided a more complex representation of understandings than is afforded by the literature, or at least, participants' understandings are not transcontextual (Behrens & Smith, 1996). I propose that investigations into teaching and learning afford a more equal balance between the influence of individual dispositions *and* the influence of contexts.

The third subsidiary research question was to determine how the complexity of people's understandings is best represented, as dichotomies, dimensions, hierarchies, or networks. There is little evidence in participants' transcripts for *dichotomies*. I have portrayed understandings as Domains, Constructs, Themes and Elaborations in the many Figures included herein. This classification system is loosely hierarchical, but that hierarchy has been imposed by me in the process of bringing some order to the data. I have questioned the *dimensions* of surface and deep approaches to studying, and the *hierarchical* conceptions of learning, and have posited that such schemes fail to take into account interactions between numerous variables. My interpretation is that participants' understandings are best conceptualised as networks of understandings, containing patterns of multiple variables and multiple interconnections. This broader conceptualisation seems worthy of further investigation.

REFERENCES

Alagumalai, S. (1999) *Reflection and translation in physics problem solving using artificial intelligence as a monitoring tool.* Unpublished PhD Thesis, Flinders University of South Australia, Adelaide, South Australia.

Anderson, J. R. (1987) Methodologies for studying human knowledge. *Behavior and Brain Sciences, 10*(3).

Anderson, J. R. (2000) *Cognitive Psychology and its Implications*. (5th ed.). San Francisco: W. H. Freeman.

Anderson, J. R., Greeno, J. G., Reder, L. M., & Simon, H. A. (2000) Perspectives on learning, thinking and activity. *Educational Researcher, 29*, 11-13.

Baird, J. R., Fensham, P. J., Gunstone, R. F., & White, R. T. (1991) The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching, 28*(2), 163-182.

Bandura, A. (1997) *Self-efficacy. The exercise of control*. New York: W. H. Freeman.

Behrens, J. T., & Smith, M. L. (1996) Data and data analysis. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 945-989). New York: Macmillan.

Bereiter, C. (2000) Keeping the brain in mind. *Australian Journal of Education, 44*(3), 226-239.

Bereiter, C., & Scardamalia, M. (1989) Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, Learning and Instruction. Essays in Honour of Robert Glaser*. (pp. 361-392). Hillsdale, NJ: Erlbaum.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001) The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 711*, 139-149.

Biggs, J. B. (1979) Individual differences in study processes and the quality of learning outcomes. *Higher Education, 8*, 381-394.

Biggs, J. B. (1987) *Student Approaches to Learning and Studying.* Burwood, Victoria: ACER.

Biggs, J. B. (1999a) *Teaching for quality learning at university*. Buckingham: Open University Press.

Biggs, J. B. (1999b) What the student does: teaching for enhanced learning. *Higher Education Research and Development, 18*(1), 57-75.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999) *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brown, A. L., & Campione, J. C. (1996) Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-270). Cambridge, Massachusetts: MIT Press.

Bruning, R. H., Schraw, G. J., & Ronning, R. R. (1999) *Cognitive psychology and instruction*. (3rd ed.). Upper Saddle River, N.J.: Merrill.

Cobb, P. (1994) Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*, 13-20.

Cobb, P. (2001) Supporting the improvement of learning and teaching in social and institutional context. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction. Twenty five years of progress* (pp. 455-478). Mahwah, NJ: Lawrence Erlbaum Associates.

Cusworth, R. A. (1995, April 18-22) *The framing of educational knowledge through "Show - and - Tell" in elementary classrooms.* Paper presented at the Annual Meeting of the American Educational Research Association.

Day, C. (1999) *Developing teachers: the challenges of lifelong learning*. Philadelphia, PA: Falmer Press.

Dweck, C. (1999) *Self Theories: Their Role in Motivation, Personality and Development.* Philadelphia, PA: Psychology Press.

Dweck, C. S. (1986) Motivational processes affecting learning. *American Psychologist, 41*(10), 1040-1048.

Edwards, D., & Mercer, N. (1993) *Common knowledge. The development of understanding in the classroom.* London: Routledge.

Eisner, E. W. (2000) Those who ignore the past...: 12 ‘easy’ lessons for the next millenium. *Journal of Curriculum Studies, 32*, 343-357.

Eklund-Myrskog, G. (1998) Students' conceptions of learning in different educational contexts. *Higher Education, 35*, 299-316.

Entwistle, N., Hanley, M., & Hounsell, D. (1979) Identifying Distinctive Approaches to Studying. *Higher Education, 8*, 365-380.

Entwistle, N., Skinner, D., Entwistle, D., & Orr, S. (2000) Conceptions and beliefs about "Good Teaching": An integration of contrasting research areas. *Higher Education Research and Development, 19*(1), 5-26.

Flavell, J. H. (1979) Metacognition and cognitive monitoring. *American Psychologist, 34*, 906-911.

Helfgott, D., & Westhaver, M. (2000) Inspiration (Version 6.0b): Inspiration Software© Inc.

Hill, L. (1999) *Making it Harder and Better: Improving Teaching and Learning*. Retrieved 24th April, 1999, from http://ultibase.rmit.edu.au/Articles/may99/hill1.htm

Hirst, P. H. (1971) What is teaching? *Journal of Curriculum Studies, 3*(1), 5-18.

Hofer, B. K. (2000) Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*, 378-405.

Hofer, B. K., & Pintrich, P. R. (1997) The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Hogan, K. (1999) Thinking aloud together: A test of an intervention to foster students collaborative scientific reasoning. *Journal of Research in Science Teaching, 36*, 1085-1109.

Jakubowski, E., & Tobin, K. (1997) Teachers' personal epistemologies and classroom learning environments. In S. Hollingsworth (Ed.), *International Action Research. A Casebook for Educational Reform*, (pp. 201-214). London: Falmer.

Karabenick, S. A. (1998) Help seeking as a strategic resource. In S. A. Karabenick (Ed.), *Strategic help seeking. Implications for learning and teaching* (pp. 1-11). Mahwah, NJ: Erlbaum.

Kerr, D. H. (1981) The structure of quality in teaching. In J. F. Soltis (Ed.), *Philosophy and education. Eightieth Yearbook of the National Society for the Study of Education* (Vol. Part I, pp. 61-93). Chicago, Illinois: University of Chicago Press.

Klein, P. D. (1996) Preservice teacher's beliefs about learning and knowledge. *Alberta Journal of Educational Research, XLII*(4), 361-377.

Koschmann, T. (2001) A third metaphor for learning: Toward a Deweyan form of transactional inquiry. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction. Twenty five years of progress* (pp. 439-454). Mahwah, NJ: Lawrence Erlbaum Associates.

Lakomski, G. (2000) Editor's introduction: Education and cognition--Naturalistic explorations into the nature of mind and learning. *Australian Journal of Education, 44*(3), 203-208.

Lampert, M. (1990) When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*, 29-63.

Lave, J. (1988) *Cognition in practice: mind, mathematics, and culture in everyday life*. Cambridge: Cambridge University Press.

Marshall, D., Summers, M., & Woolnough, B. (1999) Students' conceptions of learning in an engineering context. *Higher Education, 38*, 291-309.

Marton, F., Dall'Alba, G., & Beaty, E. (1993) Conceptions of Learning. *International Journal of Educational Research, 19*(3), 277-299.

Miles, M. B., & Huberman, A. M. (1994) *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks: Sage Publications.

Murphy, P. K., & Alexander, P. A. (2000) A Motivated Exploration of Motivation Terminology. *Contemporary Educational Psychology, 25*, 3–53.

Nelson, T. O. (1996) Consciousness and metacognition. *American Psychologist, 51*(102-116).

Newsome III, G. L. (2000) A review of some promising approaches to understanding and improving thinking skills. *Journal of Research and Development in Education, 33*(199-222).

Nuthall, G. (2000a) The role of memory in the acquisition and retention of knowledge in science and social studies units. *Cognition and Instruction, 18*(1), 83-139.

Nuthall, G. (2000b) The anatomy of memory in the classroom: Understanding how students acquire memory processes from classroom activities in science and social studies units. *American Educational Research Journal, 37,* 247-304.

Osborne, R., & Freyberg, P. (Eds.). (1980) *Learning in Science. The implications of children's science*. Auckland: Heinemann.

Paris, S. G., & Ayres, L. R. (1994) *Becoming reflective students and teachers: with portfolios and authentic assessment*. Washington, DC: American Psychological Association.

Perry, W. G. (1970) *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.

Phillips, D. C. (2000) An opinionated account of the constructivist landscape. In D. C. Phillips (Ed.), *Constructivism in education. Opinions and second opinions on controversial issues. Ninety-ninth yearbook of the national society for the study of education* (pp. 1-16). Chicago, Illinois: University of Chicago Press.

Pintrich, P. R. (2000) Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544-555.

Prosser, M., & Trigwell, K. (1997) Relations between perceptions of the teaching environment and approaches to teaching. *British Journal of Educational Psychology, 67*, 25-35.

QSR. (1997) QSR NUD*IST (Version 4): QSR International.

Saljo, R. (1979) *Learning in the learner's perspective: 1. Some common-sense conceptions* (Report 76): Gothenberg: University of Gothenberg, Department of Education.

Samuelowicz, K., & Bain, J. D. (1992) Conceptions of teaching held by academic teachers. *Higher Education, 24*, 93-111.

Schommer, M. (1990) Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82,* 498-504.

Schommer, M. (1993) Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*, 406-411.

Schommer, M., Crouse, A., & Rhodes, N. (1992) Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology, 84*, 435-443.

Schunk, D. H. (1995) Inherent details of self-regulated learning include student perceptions. *Educational Psychologist, 30*, 213-216.

Schunk, D. H., & Ertmer, P. A. (1999) Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology, 91*, 251-260.

St Julien, J. (2000) Changing conceptions of human intelligence and reasoning: Implications for the classroom. *Australian Journal of Education, 44*(3), 254-271.

Sternberg, R. J. (2000) *Practical intelligence in everyday life*. Cambridge, U.K.: Cambridge University Press.

Stigler, J., W, & Hiebert, J. (1999) *The teaching gap.* New York: The Free Press.

Sturman, A. (1999) Case study methods. In J. P. Keeves & G. Lakomski (Eds.), *Issues in educational research* (pp. 103-112). New York: Pergammon.

Tasker, R. (1981) Children's views and classroom experiences. *The Australian Science Teachers Journal, 27*(3), 33-37.

Tasker, R. (1992) Effective teaching. What can a constructivist view of learning offer? *The Australian Science Teachers Journal, 38*(1), 25-34.

Tasker, R., & Freyberg, P. (1985) Facing the mismatches in the classroom. In R. Osborne & P. Freyberg (Eds.), *Learning in Science* (pp. 66-80). Auckland: Heinemann.

Tasker, R., & Osborne, R. (1983) Portraying pupils' classroom experiences. *Research in Science & Technological Education, 1*, 133-144.

Taylor, P. G. (1996) Reflections on students' conceptions of learning and perceptions of learning environments. *Higher Education Research and Development, 15*(2), 223-237.

Tobin, K., Tippins, D. J., & Gallard, A. J. (1994) Research on instructional strategies for teaching science. In D. L. Gabel (Ed.), *Handbook of Research on Science Teaching and Learning* (pp. 45-93). New York: Macmillan.

Trigwell, K., & Prosser, M. (1996) Congruence between intention and strategy in university science teachers' approaches to teaching. *Higher Education, 32*, 77-87.

Trigwell, K., & Prosser, M. (1997) Towards an understanding of individual acts of teaching and learning. *Higher Education Research and Development, 16*(2), 241-252.

Trigwell, K., Prosser, M., & Taylor, P. (1994) Qualitative differences in approaches to teaching first year university science. *Higher Education, 27*, 75-84.

Trigwell, K., Prosser, M., & Waterhouse, F. (1999) Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education, 37*, 57-70.

von Glasersfeld, E. (1995) A constructivist approach to teaching. In L. P. Steffe & J. Gale (Eds.), *Constructivism in Education* (pp. 3-15). Hillsdale, NJ: Erlbaum.

Wenger, E. (1998) *Communities of Practice. Learning, Meaning, and Identity*. Cambridge, UK: Cambridge University Press.

Westwood, P. (1997) *Commonsense methods for children with special needs*. (3rd ed.). London: Routledge.

Whitehead, A. N. (1942) *The aims of education: and other essays*. London: Williams & Norgate.

Wigfield, A., & Eccles, J. S. (2000) Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.

Wilson, E. O. (1998) *Consilience*. New York: Alfred A. Knopf, Inc.

Winne, P. H. (1995) Inherent details in self-regulated learning. *Educational Psychologist, 30*, 173-187.

Winne, P. H., & Marx, R. W. (1982) Students' and teachers' views of thinking processes for classroom learning. *The Elementary School Journal, 82*, 493-518.

Zimmerman, B. J. (1998) Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self reflective practice* (pp. 1-19). New York: Guilford Press.

Zimmerman, B. J., & Kitsantsis, A. (1997) Developmental phases in self regulation: shifting from process goals to outcome goals. *Journal of Educational Psychology, 89*, 29-36.

International Education Journal Vol 2, No 4, 2001
Educational Research Conference 2001 Special Issue
http://www.flinders.edu.au/education/iej

Students' Achievement and Perceptions of School Climate During the Transition from Single Sex Education to Co-education

Shirley M. Yates
Flinders University, School of Education

In 1999, a non-government school with a long tradition of "boys only" education in South Australia introduced co-education at the secondary school level. The following year, girls were admitted at the primary school level. Educational achievement and perceptions of the psychosocial climate of the school's learning environment were measured in all primary and secondary students at the end of this momentous year and again one year later. Questionnaires measuring students' perceptions of the relationship dimensions of cohesiveness, friction and satisfaction and the personal dimensions of competitiveness and difficulty were administered, together with a test of achievement. Interesting differences were found in student achievement and perceptions of the school's climate, over time and across the grade levels.

Achievement; Perceptions; School Climate; Co-education.

INTRODUCTION

This longitudinal study investigated students' educational achievement and perceptions of the school learning environment during the two-year transition from single sex education to co-education. In 1999, a non-government school in metropolitan Adelaide, South Australia with a 151-year tradition of single sex education for boys, began to offer education to both girls and boys. As the school had been established in 1848 to provide single sex education for primary and secondary school aged boys, the decision to introduce co-education at all year levels was an important innovation. In the first year of the introduction of co-education, girls were enrolled in the Middle School Years 7 to 10 and Senior Years 11 to 12, and in the following year girls were admitted for the first time to the Primary School Years 3 to 6.

The study was conducted during a period of considerable institutional change within the school. In the same year that co-education was introduced initially many other changes were effected. A new school Principal was appointed, new teachers were employed and a Middle School, embracing Years 7 to 10, was formed. A director was engaged to co-ordinate this Middle School. In addition, a building refurbishment and expansion program was initiated.

Single Sex and Co-education

Over the past three decades the relative merits of single sex and co-education for the educational and socio-emotional development of school-aged students, particularly at the secondary school level, have been debated extensively. Some research evidence has been supportive of co-education, while other studies have cited the benefits of single sex education (see, Woodward, Fergusson & Horwood, 1999). In relation to academic achievement, studies of the effects of single sex and co-education at the secondary level have yielded inconsistent results. In general, some studies provide support for the benefits of co-education, (Marsh, 1989, Marsh, Smith, Marsh & Owens, 1988), others support single sex education (Austin,

1977, Lee & Bryk, 1986, Riordan, 1985) and yet others report no differences in achievement (Miller & Dale, 1974, Rutter, Maughan, Mortimer & Ouston, 1979).

The Transition from Single Sex to Co-education

Relatively few studies have examined the effects of the transition from single sex education to co-education on student outcomes. In the mid 1980s in Sydney, Australia, a non-government boys' high school and a non-government girls' high school were reorganized into two co-educational schools. Longitudinal information was collected before, during and after the transition to co-education (Marsh, 1989, Marsh *et al.,* 1988). No significant differences in achievement in Mathematics or English were found across the five years of the study. However, while there was a small decrease in self-concepts for students attending the co-educational classes in the transitional year, there was an overall increase in multidimensional self-concepts before and after the introduction of co-education.

In a similar study conducted between 1985 and 1987 in the United States of America, a single sex girls' secondary school merged with a similar boys' school to form a mixed sex school, located on the two campuses of the two original schools. Girls' attitudes towards mathematics were measured in the single sex situation, with both male and female students studying mathematics surveyed following the merger (Steinbeck & Gwizdala, 1995). Traditional gender differences were found in relation to self-confidence, perceptions of the usefulness of mathematics and classroom behaviour. While the attitudes of the girls from the single sex school remained positive towards mathematics after the merger, their predictions that there would be changes in the climate of the classroom were borne out. Not only did the girls report feeling intimidated, hesitant, uncomfortable and dumb during mathematics lessons in the mixed sex classrooms, they also perceived that teachers gave more attention to the boys (Steinbeck & Gwizdala, 1995).

In 1996 a single sex non-government boys' school in Adelaide, South Australia became co-educational. Academic achievement, optimism, pessimism and overall explanatory style were measured in all primary and secondary boys in Years 3 to 10 in the last year of single sex education, during the transitional year and one year after the introduction of co-education (Yates, 2000). In the first year of the study, data were also collected from all primary and secondary girls in Years 3 to 10 in a comparable single sex girls' school. By comparison with the girls, the boys in the single sex school had higher scores on the Word Knowledge tests (Thorndike, 1973) used as the general measure of educational achievement. These gender differences were significant at Year 3 at the primary school level and the secondary levels of Years 8, 9 and 10. In both the year of transition and the year following the introduction of co-education, increases were evident in the boys' achievement, particularly in Years 4, 5, 9, 10, 11 and 12. There was also a significant decrease in optimism and in total explanatory style associated with the transition to co-education, particularly for boys in Years 8 and 9, but this trend was not evident one year later.

School Climate

The climate of a school has an important effect on student learning (Fraser, 1994). While much of the research into school climate has focused on teachers' perceptions (Fraser, 1994), many studies have indicated that students' perceptions of the psychosocial aspects of the learning environment of their classroom are clearly related to motivational and achievement outcomes (Fraser, 1998). A meta-analysis of studies involving 17,805 students in 823 classes in eight subject areas across four nations revealed that student achievement was enhanced in classrooms with greater Cohesiveness, Satisfaction and Goal Direction and less Disorganization and Friction (Haertel, Walberg & Haertel, 1981). The concepts of Cohesiveness, Satisfaction, Goal Direction, Disorganization and Friction used in the measurement of classroom climate have their origins in Moos' (1974) scheme for classifying human environments. Cohesiveness, Satisfaction and Friction scales are encompassed within

the Relationship Dimensions (Moos, 1974) whereas Goal Direction and Disorganization are part of the System Maintenance and Change Dimensions that include scales of Difficulty and Competitiveness. While these three dimensions have been studied in many different environments (Fraser, 1998), they have not been measured at the school level during the transition from single sex education to co-education.

AIMS

The aims of this study were to:

1. measure student educational achievement and perceptions of school climate during the transition from single sex education to co-education;
2. monitor and evaluate changes in students' educational achievement and their perceptions of the climate of the school over time; and
3. examine Year level differences in achievement and perceptions.

METHOD

Participants

In 1999, the study involved all students in Years 3 to 11 in the school (N = 484), with girls being present only in Years 7 to 11. In 2000, all students in Years 3 to 12 participated (N = 597). The numbers of boys and girls in Years 3 to 12 in 1999 and 2000 who participated in the study are presented in Table 1.

Table 1. Number of students by Year level and gender in 1999 and 2000

1999	**Yr3**	**Yr4**	**Yr5**	**Yr6**	**Yr7**	**Yr8**	**Yr9**	**Yr10**	**Yr11**	**Yr12**	**Total**
Boys	25	31	34	40	40	49	71	75	78		441
Girls					9	2	3	16	13		43
Total					**49**	**51**	**74**	**91**	**91**		**484**
2000											
Boys	19	33	31	37	45	54	52	76	78	80	505
Girls	5	3	7	9	6	17	6	4	20	15	92
Total	**24**	**36**	**38**	**45**	**51**	**71**	**58**	**80**	**98**	**95**	**597**

Instruments

Word Knowledge Tests

A general measure of educational achievement was obtained from the *Word Knowledge Test 1, 2 or 3* (Thorndike, 1973). These tests have been used previously in international studies as a general assessment of students' underlying achievement and scholastic aptitude. Each test consists of 40 word pairs, with the student asked to rate whether the words are the same or opposite in meaning. Thirteen word pairs are common to Test 1 and Test 2 and twenty items are common to Test 2 and Test 3. One word pair is common to all three tests. Sample items from Test 1, Test 2 and Test 3 are presented in Table 2.

School Climate

Primary school students' perceptions of the climate of the school were measured with *My School Inventory*, while the *School Learning Environment Inventory* was used to ascertain the perceptions of the students at the secondary school level. *My School Inventory* was adapted from *My Class Inventory* (see, Fisher and Fraser, 1981; Fraser, Anderson and Walberg, 1982) and the *School Learning Environment Inventory* had its origins in the *Learning Environment*

Inventory (Anderson & Walberg, 1974); Fraser, Anderson and Walberg, 1982). *My School Inventory* and the *School Learning Environment Inventory* contained five common sub-scales measuring students' perceptions of the Relationship and Personal Development Dimensions (Moos, 1974) of the school during the transition to co-education. The Relationships Dimensions, measured through three sub-scales of Cohesiveness, Friction and Satisfaction, tapped the nature and intensity of personal relationships, the extent of student involvement and the degree of mutual support and help within the school learning environment. The Personal Development Dimensions assessed in two subs-scales of Competitiveness and Difficulty, measured students' personal growth and self-enhancement during the transitional process. Sample items from the Relationship and Personal Development Dimensions within *My School Inventory* (MSI) and *School Learning Environment Inventory* (SLEI) are presented in Table 3.

Table 2. Sample items from Word Knowledge Test 1, Test 2 and Test 3

Word Knowledge Test 1		
1	quick	slow
40	scarce	rare
Word Knowledge Test 2		
1	savoury	insipid
40	obvious	indisputable
Word Knowledge Test 3		
1	acquire	dispel
40	ephemeral	eternal

Table 3. Sample items for the Relationship and Personal Development Dimensions within My School Inventory (MSI) and School Learning Environment Inventory (SLEI)

Relationship Dimensions:		
Cohesiveness		
MSI	Item 28	Children in our school like each other as friends
SLEI	Item 31	Students in the school are frequently personal friends
Friction		
MSI	Item 2	Children are always fighting with each other
SLEI	Item 6	There is constant bickering among student
Satisfaction		
MSI	Item 16	Most children say the school is fun
SLEI	Item 73	Students look forward to coming to classes
Personal Development Dimensions		
Competitiveness		
MSI	Item 24	Some students try to do their work better than the others
SLEI	Item 30	Students compete to see who can do the best work
Difficulty		
MSI	Item 3	In our school, work is hard to do
SLEI	Item 54	Students tend to find the schoolwork hard to do

Procedure

The test of educational achievement and the questionnaire measuring perceptions of the school climate were administered to all students in their classrooms in October, 1999 and October, 2000 at the same time and on the same day across the school.

Academic achievement: Students in Years 3 to 7 were administered *Word Knowledge Test 1* (Thorndike, 1973), those in Years 8 to 10, *Word Knowledge Test 2* (Thorndike, 1973) while *Word Knowledge Test 3* (Thorndike, 1973) was completed by boys and girls in Years 11 and 12.

School Climate: Primary school students in Years 3 to 7 were administered *My School Inventory*, with the *School Learning Environment Inventory* given to the secondary level students in Years 8 to 12. To provide a common group for equating purposes students in Years 8 and 9 also completed *My School Inventory*.

Analysis

The validity of the word knowledge tests and the school climate questionnaires were examined with QUEST (Adams and Khoo, 1994) and all non-fitting items deleted. The Rasch scaling procedure (Rasch, 1960; 1966) was used to calibrate the word knowledge tests and questionnaires to bring them to common interval scales. A single *Word Knowledge* scale of achievement was formed using the common item linking procedure, with the calibration based on students who answered all items. *School Environment* Scales were formed for the Relationship Dimensions sub-scales of Cohesiveness, Friction, Satisfaction and the Personal Development Dimensions sub-scales of Competitiveness and Difficulty, found in both the *My School Inventory* and *School Learning Environment Inventory*, using the common persons linking procedure. Case scores were then estimated concurrently for the *Word Knowledge* Scale and five sub-scale dimensions of the *School Environment* for all students from Years 3 to 11 for 1999 and Years 3 to 12 for 2000.

RESULTS

The results of the univariate analyses of variance for the *Word Knowledge* Scale and five sub-scales of *School Environment*, undertaken with the case estimate scores, are presented in Table 4. Differences between the means were examined across the Year levels and over time. In view of the relatively small numbers of girls in both 1999 and 2000, gender differences are not examined. Significant differences were found across the Year levels for Word Knowledge, Cohesiveness, Friction, Satisfaction and Difficulty, across time for Friction, Satisfaction, and Competitiveness and between the boys and girls for Cohesiveness only. It should be noted that when testing for statistical significance the student, rather than the class was used as the unit of analysis. While students at the primary school level are grouped in classes within the school, students at the secondary level change classes according to the subjects that they are studying. It was necessary therefore to focus the analysis at the student level, but it is recognized that this assumption may not be completely appropriate at the primary school level, where the class is probably the operational unit.

Table 4. Analysis of Variance: Word Knowledge and School Environment Sub-Scales by Year level and Time

Variable	**Year**	**Time**
Word Knowledge	F = 24.62***	NS
Cohesiveness	F = 4.91***	NS
Friction	F = 23.60***	F = 10.06***
Satisfaction	F = 59.45***	F = 38.88***
Competitiveness	NS	F = 3.83*
Difficulty	F = 26.20***	NS

***p<.001, **p<.01, * p<.05, NS= Not Significant

Note: F based on 262 degrees of freedom for the error item.

Educational Achievement

The results of the *Word Knowledge* Scale for all students in 1999 and 2000 are presented in Figure 1.

In both 1999 and 2000 there is an overall significant trend of increasing achievement across the Year levels ($F = 24.62$, $p < .001$). There is a marked improvement in achievement for students in Years 5 and 6 in 2000, with a slight drop in Year 8 for the same year. Relative to 1999 there is also an increase in achievement in Years 9 and 10 in 2000 and from Year 11 in 1999 to Year 12 in 2000. There is a slight decrease in Year 11 over the same period.

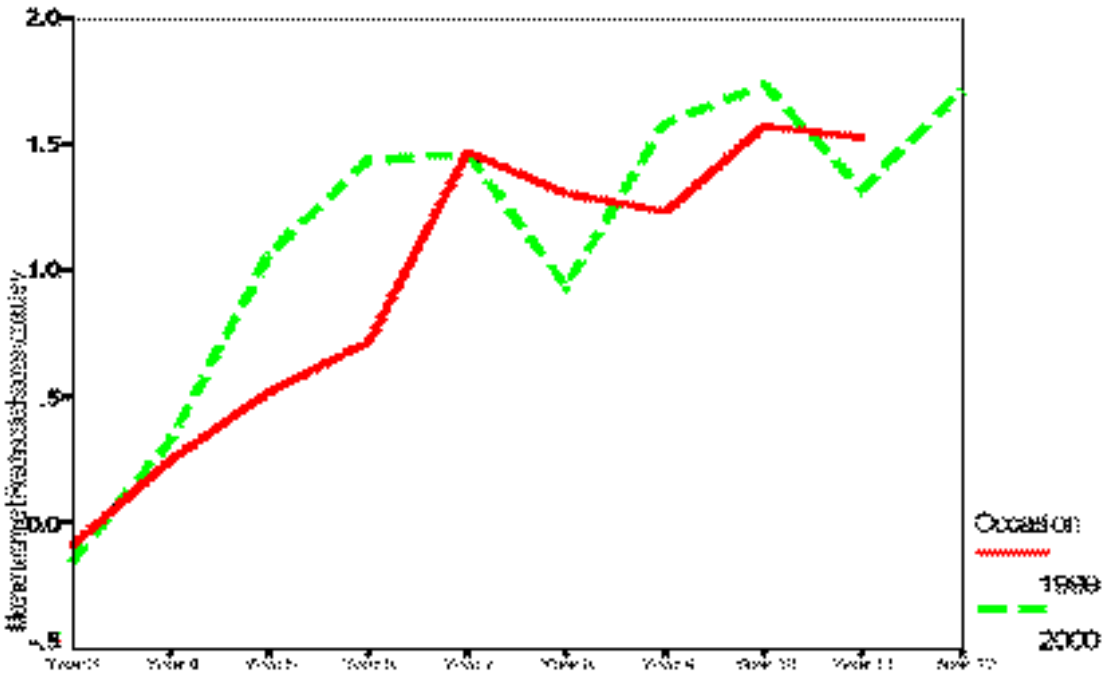


Figure 1: Word Knowledge scores for 1999 and 2000

School Learning Environment

Results of students' perceptions of the school learning environment in 1999 and 2000 are presented separately within the three Relationship Dimensions of Cohesiveness, Friction and Satisfaction and the two Personal Development Dimensions of Competitiveness and Difficulty.

Relationship Dimensions:

The Relationship Dimension of Cohesiveness

The student cohesiveness sub-scale measures the extent to which students know, help and are supportive of one another within the school. The sub-scale was formed from seven *My School Inventory* items and seven items from the *School Learning Environment Inventory*. The results for all students from Years 3 to 11 in 1999 and Years 3 to 12 in 2000 are presented in Figure 2.

In Figure 2 there is a noticeable and significant improvement in cohesiveness across the school from 1999 to 2000 ($F = 4.81$, $p < .001$). Students entering Year 3 in 2000 report very high levels of cohesion. There is also a high level of cohesion in Year 6 in 2000. Students from Year 9 who have the lowest level of cohesiveness in 1999 report much higher levels of cohesion in

Year 10 in 2000. Students in Year 12 in 2000 also give evidence of slightly more cohesiveness than the previous year.

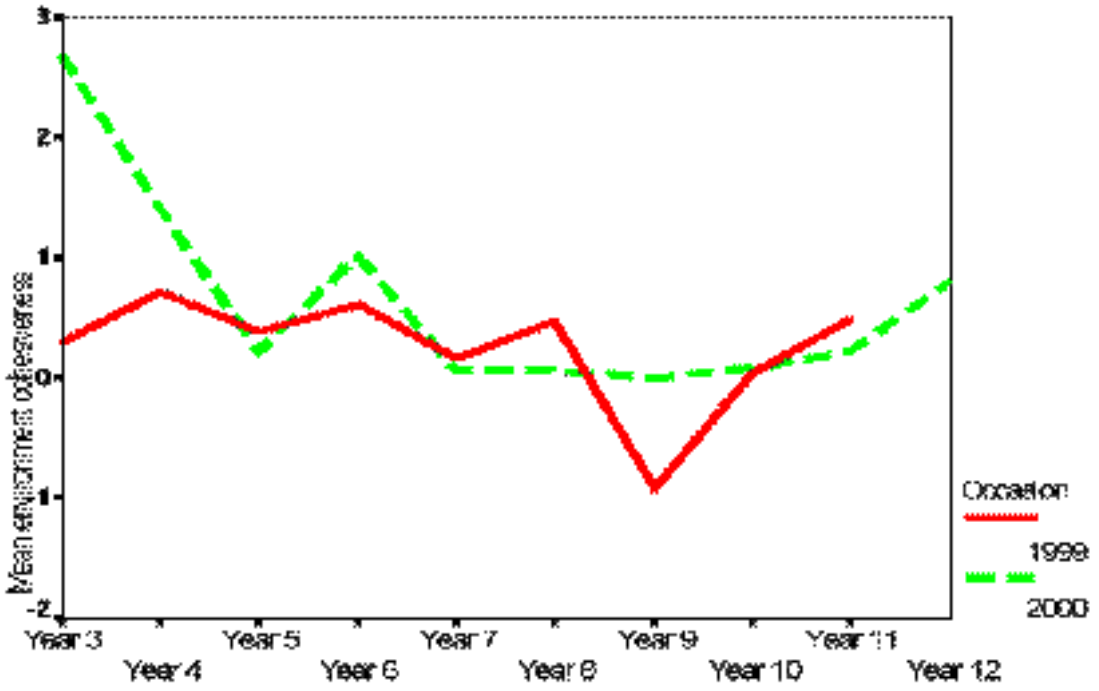


Figure 2: Relationship Dimension: Cohesiveness scores 1999 and 2000

The Relationship Dimension of Friction

This dimension is concerned with the extent to which students are uncooperative, mean, argumentative and are in conflict with each other within the school. *My School Inventory* contains eight items to measure perceptions of friction, while the *School Learning Environment Inventory* has 7 Friction items. The results for the Friction sub-scale for all students in 1999 and 2000 are presented in Figure 3.

The results show some very interesting and significant differences in the pattern of students' perceptions of friction across the school ($F = 23.60$, $p < .001$) in both years of the study. ($F = 10.06$, $p < .001$). The trend from 1999 to 2000 is for an overall decrease in perceptions of friction across the school. This is particularly evident in 2000 in Years 4, 7, 9, 10 and 12. Students at the primary school level have comparatively low levels of friction, particularly in Year 3 in 2000. There is an increase in the perception of friction in Year 7 and Year 10 in 1999. The pattern for the Year 7 students continues on into Year 8 in 2000 but changes for the Year 10 students as they move into Year 11 in 2000.

The Relationship Dimension of Satisfaction

This dimension measures students' overall perceptions of happiness, pleasure and contentment with the school environment. There are nine items in *My School Inventory* and seven items in the *School Learning Environment Inventory* measuring satisfaction.

The results for the relationship dimension of satisfaction for all students for 1999 and 2000 are presented in Figure 4.

By far the biggest and most significant change in perception of the school climate from 1999 to 2000 is reflected in the dimension of satisfaction. With the exception of Year 5 where there is a slight decrease and Year 10 which does not change over time, students overall clearly report higher levels of satisfaction with the school in 2000 than they do in 1999 ($F = 38.88$, p

< .001). Interesting and very significant trends are evident across the year levels in both years ($F = 59.45$, $p > .001$). Overall students at the primary school level are very satisfied, but particularly so in Year 3 and Year 6 in 2000. Satisfaction decreases across the Middle School Years 7 to 10, although this is much less so in 2000 than it is in 1999. The negative levels of satisfaction, clearly evident in Year 11 in 1999 improve in Year 12 in 2000, as is the case from Year 10 in 1999 to Year 11 in 2000.

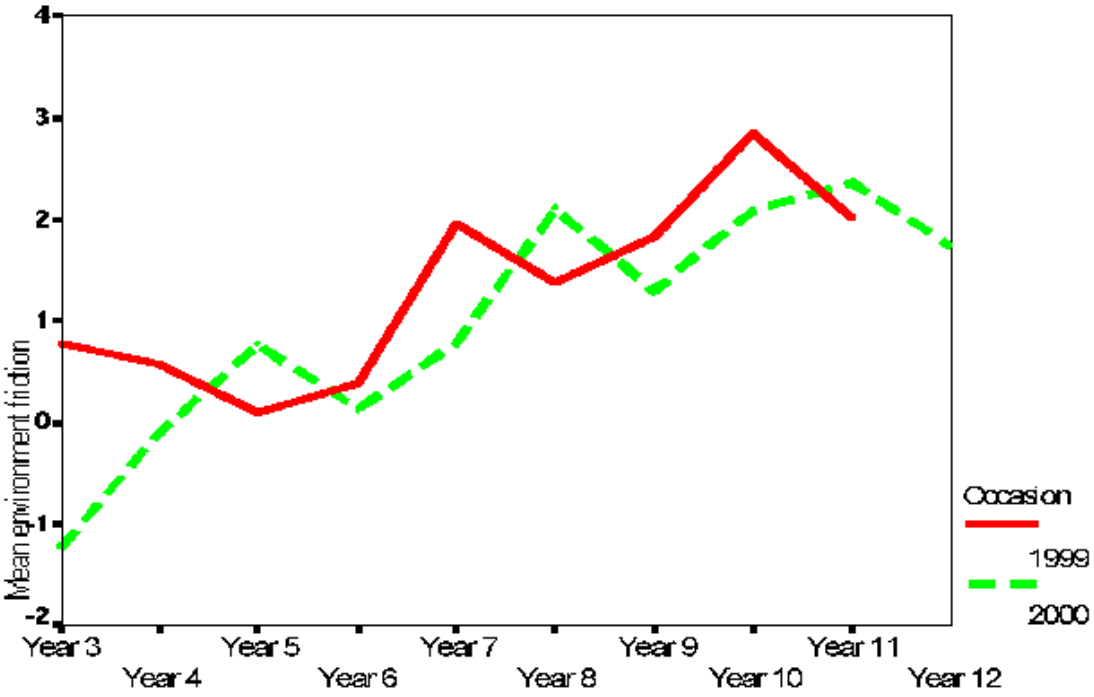


Figure 3: Relationship Dimension: Friction scores 1999 and 2000

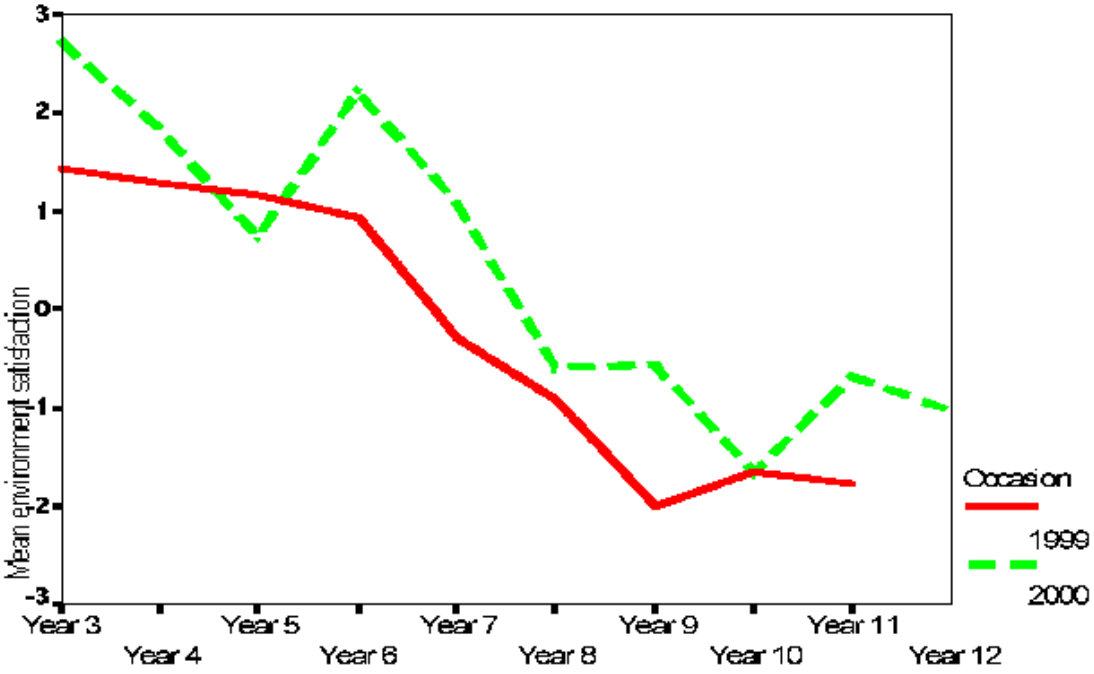


Figure 4: Relationship Dimension: Satisfaction scores 1999 and 2000

Personal Development Dimensions

The Personal Development Dimension of Competitiveness

The competitive dimension of personal development is concerned with the extent to which students are ambitious, keen to outdo each other and view the other as rival contestants. The sub-scale is formed from seven items in *My School Inventory* and seven items in the *School Learning Environment Inventory*. The results of the competitiveness dimension for 1999 and 2000 across the school are presented in Figure 5.

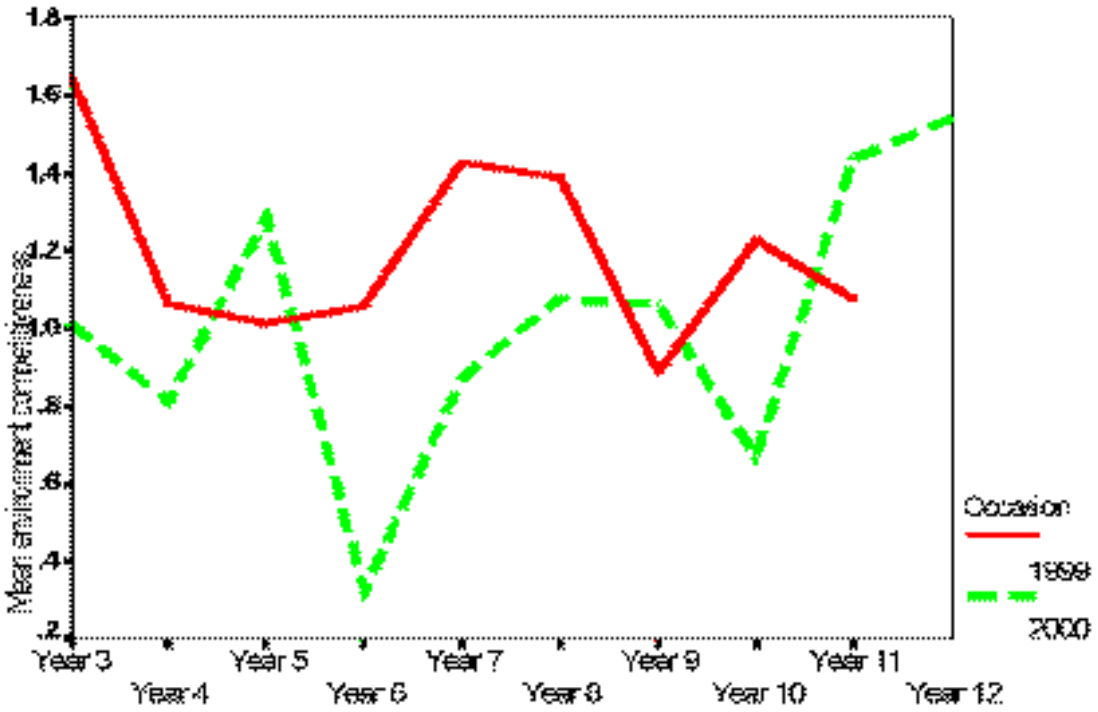


Figure 5: Personal Dimension: Competitiveness scores 1999 and 2000

Figure 5 shows a very interesting pattern of competitiveness across the school. Overall, there is a significant decrease in competitiveness over time ($F = 3.83$, $p > .05$), particularly in Years 3, 4, 6, 7, 8 and 10 in 2000. Contrary to this general pattern, there is an increase in competitiveness in Years 5, 11 and 12 in 2000. In 1999, highest levels are reported in Years 3, 7 and 8, and in 2000 in Years 5, 11 and 12. Students in Year 6 report the lowest levels of competitiveness in 2000.

The Personal Development Dimension of Difficulty

The items in both *My School Inventory* and the *School Learning Environment Inventory* for this dimension are concerned with the extent to which students perceive schoolwork to be demanding, complicated and troublesome. *My School Inventory* contains eight items measuring difficulty while the *School Learning Environment Inventory* has seven items for this dimension.

The results for the difficulty dimension across Years 3 to 11 in the school in 1999 and Years 3 to 12 in 2000 are presented in Figure 6.

Overall, in both 1999 and 2000, students do not perceive their schoolwork to be difficult. There are significant differences in perception of difficulty between students in the primary school years and those in the Middle and Senior years ($F = 26.20$, $p < .001$) in both 1999 and 2000, with Senior students in Years 11 and 12 only reporting levels of difficulty above the mean.

Year 3 students report work to be more difficult than do their counterparts in Years 4, 5 and 6 in both 1999 and 2000, although their perceptions of difficulty are less so in 2000. The biggest change in the perception of schoolwork difficulty is with students in Years 7 and 8 in 2000. Students in Year 6 in both 1999 and 2000 and Year 5 in 2000 only express the lowest levels of perception of difficulty across the school.

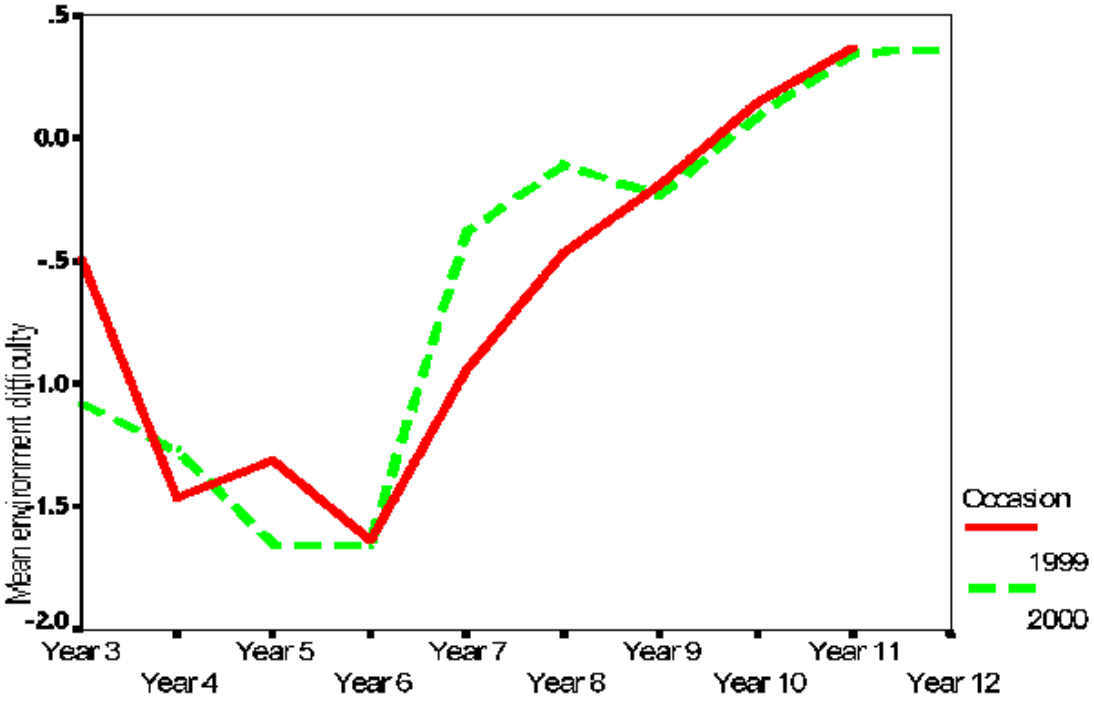


Figure 6: Personal Dimension: Difficulty scores 1999 and 2000

Summary of Results

In the transition to co-education the following trends were evident across the school:

1. a steady and significant pattern of growth in general achievement across the year levels, with marked increases evident in Years 5, 6, 10 and 12 in 2000

2. in the Relationship Dimensions of the school climate:
   increases in levels of student cohesiveness from 1999 to 2000
   significant reduction in perceptions of friction in 2000, and
   greatest improvement evident in students' satisfaction with their school life

   However, the pattern of increasing friction and decreasing satisfaction across the year levels does not change over time.

3. in the Personal Development Dimensions of the school climate:
   a marked decrease in competitiveness between students in 2000, and
   a decrease in perceptions of the difficulty of schoolwork in Years 3 to 10 in 2000, although there is a pattern of increasing difficulty of schoolwork across the year levels in both years.

DISCUSSION AND CONCLUSIONS

The information presented in this study is unique because it captures both primary and secondary students' perceptions across the school over time, in a climate of considerable institutional change. The increases in educational achievement and the positive changes in the Relationship Dimensions of Cohesiveness, Friction and Satisfaction and Personal Development Dimensions of Competitiveness and Difficulty would indicate that the introduction of co-education and associated structural changes in the school had positive benefits for student outcomes.

These results lend support to the benefits of co-education for student achievement cited by Marsh (1989) and Marsh *et al*., (1988). The findings also suggest that the relationship between the Relationship and Personal Dimensions and student achievement found in a very large number of studies (Haertel, Walberg & Haertel, 1981), may operate not only at the level of the climate of the classroom, but also at the school level. Clearly, primary and secondary students in this school have responded positively to the change from single sex education to co-education, both in terms of their achievement and their perceptions of the psychosocial climate.

It should be noted that this study has been conducted only over the two year period within which co-education was phased into the school. In two of the three studies of the transition from single sex to co-education (Marsh *et al.*, 1988; Steinbeck & Gwizdala, 1995; Yates, 2000) negative changes were evident in student characteristics that were not maintained once the transitional period had passed (Marsh *et al.,* 1988; Yates, 2000). The conduct of the third study (Steinbeck and Gwizdala, 1995) prior to and during the merger of two single sex campuses precludes investigations of whether girls' negative perceptions of the climate of the classroom were maintained over time. In light of these transitional differences, it is therefore planned to continue with this study annually until 2003 to ascertain whether the trends evident within the transitional period are maintained over time. It is also anticipated that over this time the proportion of girls in the school will increase.

REFERENCES

Adams, R..J & Khoo, S.K. (1993) *Quest - The interactive test analysis system.* Hawthorn, Victoria: ACER.

Anderson, G.J. & Walberg, H.J. (1974) Learning Environments. In H. J Walberg (Eds) *Evaluating educational performance: a sourcebook of methods, instruments, and examples.* Berkeley: McCutchan.

Astin, A.W. (1977) On the failure of educational policy. *Change,* 9, 40-45.

Fisher, D.L. & Fraser, B.J. (1981) Validity and use of My Class Inventory. *Science Education,* 65, 145-156.

Fraser, B.J. (1998) Science learning environments: Assessment, effects and determinants. In B. J. Fraser and K. G. Tobin (Eds), *The international handbook of science* education (pp.527-564). Dordrecht, The Netherlands: Kluwer.

Fraser, B.J., Anderson, G.J. & Walberg, H.J. (1982) *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (third version) Perth: Western Australian Institute of Technology.

Haertel, C.D., Walberg, H.J. & Haertel, E.H. (1981) Socio-psychological environments and learning: a quantitative synthesis. *British Educational Research Journal,* 7, 27-56.

Irving, J. (1981) Coeducational or single-sex schools? A review of the literature. *Set* 76, No. 1, Item 9.

Lee, V. E. & Bryk, A. S. (1986) Effects of single-sex secondary schools on student achievement and attitudes, *Journal of Educational Psychology*,78, 381-395.

Marsh, H.W., Smith, I.D., Marsh, M. & Owens, L. (1988) The transition from single-sex to coeducational high schools: Effects on multiple dimensions of self-concepts and on academic achievement. *American Educational Research Journal*, 25, 237-269.

Miller, P., McC & Dale, R. R. (1974) A comparison of the degree results of university students from coeducational and single-sex schools. *British Journal of Educational Psychology*, 44, 307-308.

Moos, R.H. (1974) *The social climate scales: An overview.* Palo Alto, California: Consulting Psychologists Press.

Rasch, G. (1960) *Probalistic models for some intelligence and attainment tests.* Copenhagen: Danmarks Paedagogiske Institut. (Republished in 1980, Chicago: University of Chicago Press).

Rasch, G. (1966) An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, 19, 49-57.

Rutter, M., Maughan, B., Mortimor, P. & Ouston, J. (1979) *Fifteen thousand hours: Secondary schools and their effects on children*, London: Open Books.

Steinbeck, M. & Gwizdala, J. (1995) Gender differences in mathematics attitudes of secondary school students. *School Science and Mathematics.* 95, (1), 36-42.

Thorndike, R.L. (1973) *Reading comprehension in fifteen countries.* New York, Wiley: and Stockholm: Almqvist and Wiksell.

Woodward, L.J., Fergusson, D.M. & Horwood, L. J. (1999) Effects of single-sex and coeducational schooling on children's academic achievement. A*ustralian Journal of Education*, 43, (2), 142-156.

Yates, S.M. (2000) Student optimism, pessimism and motivation during the transition to co-education. Paper presented at the Annual Association for Research in Education, Sydney. www.aare.edu.au/00pap/yat0439.htm

Acknowledgements

1. Professor John Keeves for his invaluable assistance with the data analyses.
2. John Harris for his exemplary research assistance.
3. Dianne Harris for data entry.
4. The Principal, staff and students of the schools who made the study possible.
5. The research was supported by a Small Australian Research Council Grant.

Changes in students' achievement in learning the Chinese Language across grades and over time

Dr Ruilan Yuan
Flinders University School of Education *Rlan.Yuan@flinders.edu.au*

This paper examines the changes in students' achievement in learning Chinese as a foreign language across school grades and over a period of a school year in an Australian school. The Rasch model using the QUEST computer program (Adams & Khoo, 1993) was employed to calculate appropriate scores to estimate the difficulty levels of the test items on a scale that operated across grade levels (Year 4 to Year 12) and across four school term occasions. This paper also identifies whether the level of students' achievement in learning the Chinese language is associated with their proficiency in English word knowledge, as well as the student's underlying verbal ability in English.

Changes, students achievement, Chinese Language

INTRODUCTION

The Chinese language program has been introduced into Australian school systems since the 1960s. Several research studies have noted factors influencing students' continuing with the learning of the Chinese language as a school subject in Australian schools (Murray & Lundberg, 1976; Fairbank & Pegalo, 1983; Tuffin & Wilson, 1990; Smith, Chin, Louie, & Mackerras, 1993). Although various reasons are reported to influence continuing with the study of the Chinese language, such as attitudinal roles, peer and family pressure, gender difference, lack of interest in languages, the measurement of students' achievement growth in learning the Chinese language across grade levels and over time, and the investigation of factors influencing such achievement growth has not been attempted. Indeed, measuring students' achievement across grade levels and over time is important in the teaching of the Chinese language in Australian school systems because it may provide greater understanding of the actual learning that occurs and enable comparisons to be made between grade levels and the teaching methods employed at different grade levels.

In the present study, in order to measure students' achievement in learning the Chinese language across grades and over time, a series of Chinese tests was designed and administered to each grade from Year 4 to Year 12 as well as over school terms from Term 1 to Term 4 in the 1999 school year. It was necessary to examine carefully the characteristics of the test items before the test scores for each student who participated in the study could be calculated in appropriate ways, because meaningful scores were essential for the subsequent analyses of the data collected in the study.

It is necessary to note that the English word knowledge tests were administered to students who participated in the study in order to examine whether the level of achievement in learning the Chinese language was associated with the proficiency in English word knowledge and the student's underlying verbal ability in English (Thorndike, 1973).

This paper comprises four sections. The first section presents the methods used for the calculation of scores. The second section considers the equating procedures, while the third

section examines the differences in scores between grade levels, and between term occasions. The last section summarises the findings from the examination of the Chinese achievement tests and English word knowledge tests obtained from the Rasch scaling procedures.

PURPOSE OF THE STUDY

The main purposes of the study are to measure student achievement in learning the Chinese language across grade levels (Year 4 to Year 12) over a time period of one school year, and to examine the relationship between students' English verbal ability and their achievement in learning the Chinese language.

THE SAMPLE

One school in South Australia was selected for this study because it met all the requirements of the study. Therefore, the subjects chosen for this study were all the students (N=945 in 39 classes) from Year 4 to Year 12 who were learning Chinese as a foreign language in the school in 1999.

Table 1. Number of Students Learning Chinese in 1999

Level	***No. of Classes***	***No. of Students***
Year 4	6	181
Year 5	6	179
Year 6	6	184
Level 1(Year 7-8)	6	175
Level 2(Year 8-9)	5	120
Level 3(Year 9-10)	5	56
Level 4(Year 10-11)	3	23
Level 5(Year 11-12)	2	27
Total	**39**	**945**

Among the 945 students who learned the Chinese language, 454 were male and 491 female. Table 5.1 presents the number of students at each grade level in 1999 who were learning the Chinese language in the school where the data were collected.

THE INSTRUMENTS

The instruments employed for collecting the data included:

1. Chinese Term 1 Test
2. Chinese Term 2 Test
3. Chinese Term 3 Test
4. Chinese Term 4 Test, and
5. Four Chinese common item tests
6. English Word Knowledge tests

RESEARCH QUESTIONS

1. Do differences exist in the level of student achievement in learning the Chinese language between grade levels?
2. Do differences exist in the level of student achievement in learning the Chinese language within grade levels across the four terms?

3. Do changes occur in level of student performance on the English word knowledge tests between grade levels?
4. Is the level of student achievement in learning the Chinese language related to performance on the English word knowledge tests?

METHODS EMPLOYED FOR THE CALCULATION OF SCORES

It was considered desirable to use the Rasch measurement procedures in this study in order to calculate scores and to provide an appropriate data set for subsequent analyses through the equating of the Chinese achievement tests, English Word Knowledge tests and attitude scales across grades and over time. In this way it would be possible to generate the outcome measures for the subsequent analyses using multivariable and multilevel analysis procedures. Lietz (1995) has argued that the calculation of scores using the Rasch model makes it possible to increase the homogeneity of the scales across grades and over occasions so that scoring bias can be minimised.

Use of Rasch Scaling

The Rasch analyses were employed in this study to measure (a) the Chinese language achievement of students across eight grades and over four term occasions; and (b) English Word Knowledge tests across grades. The estimation of the scores received from these data sets using the Rasch model involved two different procedures, namely, calibration and scoring, which are discussed below.

The raw scores on a test for each student were obtained by adding the number of points received for correct answers to each individual item in the test, and were entered into the SPSS file. With respect to the context of the current study, the calculation of these raw scores did not permit the different tests to be readily equated. In addition, the performance levels of the students and the difficulty levels of the items were not estimated on an interval scale. Hence, the Rasch model was employed to calculate appropriate scores to estimate accurately the difficulty levels of the items on a scale that operated across grade levels and across occasions. In the Chinese language achievement tests, and English Word Knowledge tests, omissions of the items were considered as wrong.

Calibration and Equating of Tests

This study used vertical equating procedures so that achievement of students in learning the Chinese language at different grade levels could be measured on the same scale. The horizontal equating approach was also employed to measure student achievement across the four term occasions. In addition, two different types of Rasch model equating, namely, anchor item equating and concurrent equating, were employed at different stages in the equating processes. The equating of the Chinese achievement tests requires common items between grades and across terms. The equating of the English Word Knowledge tests requires common items between the three tests, namely, Tests 1V, 2V and 3V. The following section reports the results of calibrating and equating of the Chinese tests between grades and across the four occasions as well as the equating of English Word Knowledge tests between grades.

Calibration and scoring of tests

There were eight grade level groups of students who participated in this study (Year 4 to Year 12). A calibration procedure was employed in this study in order to estimate the difficulty levels (i.e. threshold values) of the items in the tests, and to develop a common scale for each data set. In the calibration of the Chinese achievement test data and English word knowledge test data in this study, three decisions were made. First, the calibration was done with data for all students who participated in the study. Second, missing items or omitted items were treated as wrong in the Chinese achievement test and the English word knowledge test data in the calibration. Finally, only those items that fitted the Rasch scale were employed for calibration and scoring. This means that, in general, the items whose infit mean square values were outside

an acceptable range were deleted from the calibration and scoring process. Information on item fit estimates and individual person fit estimates is reported below.

Item fit estimates

Rasch analysis estimates the degree of fit of particular items to an underlying or latent scale, and the acceptable range of item fit taken in this study for each item in the three types of instruments, in general, was between 0.77 and 1.30. The items whose values were below 0.77 or above 1.30 were generally considered outside the acceptable range. The values of overfitting items are commonly below 0.77, while the values of misfitting items are generally over 1.30. In general, the misfitting items were excluded from the calibration analysis in this study, while in some cases it was considered necessary and desirable for overfitting items to remain in the calibrated scales.

It should be noted that the essay writing items in the Chinese achievement tests for Levels 1 and upward that were scored out of 10 marks or more were split into sub-items with scores between 0 to 5. For example, if a student scored 23 on one writing item, the extended sub-item scores for the student were 5, 5, 5, 4, and 4. The overfitting items in the Chinese achievement tests were commonly those subdivided items whose patterns of response were too predictable from the general patterns of response to other items.

Table 2 presents the results of the Rasch calibration of Chinese achievement tests across grades and over the four terms. The table shows the total number of all the items for each year level and each term, the numbers of items deleted, anchor items, bridge items and the number of items retained for analysis.

The information in Table 2 show that 17 items (6.9 per cent of the total items) did not fit the Rasch model and were removed from the Term 1 data file. There were 46 items (13 per cent) that were excluded from the Term 2 data file, while 33 items (13 per cent) were deleted from the Term 3 data file. A larger number of 68 deleted items was found in the Term 4 data file (22 per cent). This might be associated with the difficulty levels of the anchor items across terms and the bridge items between year levels because students were likely to forget what had been learned previously after learning new content. In addition, there were some items that all students who attempted these items answered correctly, and such items had to be deleted because they provided no information for the calibration analysis.

Table 2. Description of Anchor, Bridge, Deleted and Retained Items in Chinese Achievement Tests

Term Level	Number of Items for Term 1					Number of Items for Term 2					Number of Items for Term 3					Number of Items for Term 4			
	T	A	B	D	R	T	A	B	D	R	T	A	B	D	R	T	A	B	D
Year 4	12	5	5	1	11	15	5	5	-	15	18	5	12	1	17	15	5	5	1
Year 5	20	5	11	3	17	20	10	10	-	20	20	10	10	-	20	20	5	10	-
Year 6	20	5	10	3	17	20	10	10	3	17	20	10	10	-	20	20	5	10	-
Level 1	40	5	6	-	40	65	10	10	-	65	64	15	10	4	60	77	10	10	13
Level 2	60	5	10	2	58	84	10	10	4	80	49	11	9	9	41	70	6	10	21
Level 3	41	5	20	-	41	72	10	10	13	59	39	10	9	6	33	72	5	10	17
Level 4	39	5	10	5	34	52	10	10	16	36	37	10	5	13	24	51	5	5	16
Level 5	22	5	10	3	19	35	5	5	10	25	-	-	-	-	-	-	-	-	-

Notes: T = Total number of items, including anchor and bridge items; A = Anchor items, namely items across terms within the same group; B = Bridge items, namely items between year levels in the same term (e.g. In Term 1 there were 5 items between Year 4 and Year 5, 5 items between Year 5 and Year 6. Therefore, there were 11 bridge items altogether in Year 5.); D = Deleted misfitting items; R = Retained fitting items for analysis.

As a result of the removal of the misfitting items from the data files for the four terms, namely, the items outside the acceptable range of 0.77 and 1.30, 237 items for the Term 1 tests; 317 items for the Term 2 tests; 215 items for the Term 3 tests; and 257 items for the Term 4 tests fitted the Rasch scale, and were therefore retained for the four separate calibration analyses.

Table 3 provides the details of the numbers of both anchor items and bridge items that satisfied the Rasch scaling requirement after deletion of misfitting items. The figures show that 33 out of 40 anchor items fitted the Rasch model for Term 1 and were linked to the Term 2 tests. Out of 70 anchor items in Term 2, 64 anchor items were retained, among which 33 items were linked to the Term 1 tests, and 31 items were linked to the Term 3 tests. Of 58 anchor items in the Term 3 data file, 31 items were linked to the Term 2 tests, and 27 items were linked to the Term 4 tests.

Table 3. Final Number of Anchor and Bridge Items for Analysis

Level	Term 1		Term 2		Term 3		Term 4		Total
	A	B	A	B	A	B	A	B	B
Year 4	4	5	4	5	14	5	4	5	20
Year 5	5	18	10	10	10	10	5	5	43
Year 6	2	7	7	10	10	10	5	5	32
Level 1	5	6	10	10	15	10	10	5	31
Level 2	5	8	8	9	3	8	0	5	30
Level 3	5	18	8	8	4	8	1	8	42
Level 4	4	10	4	8	2	5	2	3	26
Level 5	3	10	13	4	-	-	-	-	14
Total	33	-	64	-	58	-	27	-	238

Notes: A = Anchor items B = Bridge items.

The last column in Table 3 provides the number of bridge items between grade levels for all occasions. There were 20 items for Years 4 and 5; 43 items between Years 5 and 6; 32 items between Year 6 and Level 1; 31 items between Levels 2 and 2; 30 items between Levels 2 and 3; 42 items between Levels 3 and 4; and 26 items between Levels 4 and 5.

In the analysis for the calibration and equating of the tests, the items for each term were first calibrated using concurrent equating across the grades and the threshold values of the anchor items for equating across occasions were estimated. Thus, the items from Term 1 were anchored in the calibration of the Term 2 analysis, and the items from Term 2 were anchored in the Term 3 analysis, and the items from Term 3 were anchored in the Term 4 analysis. This procedure is discussed further in a later section of this chapter.

Table 4 summarises the fit statistics of item estimates and case estimates in the process of equating the Chinese achievement tests using anchor items across the four terms. The first panel shows the summary of item estimates and item fit statistics, including infit mean square, standard deviation and infit *t*, as well as outfit mean square, standard deviation and outfit *t*. The bottom panel displays the summary of case estimates and case fit statistics as well as infit and outfit results.

Person fit estimates

Apart from the examination of item fit statistics, the Rasch model also permits the investigation of person statistics for fit to the Rasch model. The item response pattern of those persons who exhibit large outfit mean square values and *t* values should be carefully examined. If erratic behaviour was detected, those persons should be excluded from the analyses for the calibration of the items with the Rasch model (Keeves & Alagumalai, 1999).

Table 4. Summary of Fit Statistics between Terms on Chinese Tests Using Anchor Items

Statistics	Terms 1/2	Terms 2/3	Terms 3/4
Summary of item estimates & fit statistics			
Mean	0.34	1.62	1.51
SD	1.47	1.92	1.87
Reliability of estimate	0.89	0.93	0.93
Infit mean square			
Mean	1.06	1.03	1.01
SD	0.37	0.25	0.23
Outfit mean square			
Mean	1.10	1.08	1.10
SD	0.69	0.56	1.04
Summary of case estimates & fit statistics			
Mean	0.80	1.70	1.47
SD	1.71	1.79	1.81
SD (adjusted)	1.62	1.71	1.73
Reliability of estimate	0.90	0.91	0.92
Infit mean square			
Mean	1.05	1.03	1.00
SD	0.6	0.30	0.34
Infit t			
Mean	0.20	0.13	0.03
SD	1.01	1.06	1.31
Outfit mean square			
Mean	1.11	1.12	1.11
SD	0.57	0.88	1.01
Outfit t			
Mean	0.28	0.24	0.18
SD	0.81	0.84	1.08

In the data set of the Chinese achievement tests, 27 out of 945 cases were deleted from Term 3 data files because they did not fit the Rasch scale. The high level of satisfactory response from the students tested resulted from the fact that in general the tests were administered as part of the school's normal testing program, and scores assigned were clearly related to school grades awarded.

Calculation of zero and perfect scores

Zero scores received by a student on a test indicate that the student answered all the items incorrectly, while perfect scores indicate that a student answered all the items correctly. Since students with perfect scores or zero scores are considered not to provide useful information for the calibration analysis, the QUEST computer program (Adams & Khoo, 1993) does not include such cases in the calibration process.

In order to provide scores for the students with perfect or zero scores and so to calculate the mean and standard deviation for the Chinese achievement tests and the English Word Knowledge tests for each student who participated in the study, it was necessary to estimate the values of the perfect and zero scores.

In this study, the values of perfect and zero scores in the Chinese achievement and English word knowledge tests were calculated from the logit tables generated by the QUEST computer program. Afrassa (1998) used the same method to calculate the values of the perfect and zero scores of the mathematics achievement tests. The values of the perfect scores were calculated by selecting the three top raw scores close to the highest possible score. For example, if the highest raw score was 48, the three top raw scores chosen were 47, 46 and 45. After the three top raw scores were chosen, the second highest value of logit (2.66) was subtracted from the first highest logit value (3.22) to obtain the first entry (0.56). Then the third highest logit value (2.33) was subtracted from the second highest logit value (2.66) to gain the second entry

(0.33). The next step was to subtract the second entry (0.33) from the first entry (0.56) to obtain the difference between the two entries (0.23). The last step was to add the first highest logit value (3.22) and the first entry (0.56) and the difference between the two entries (0.23) so that the highest score value of 4.01 was estimated. Table 5a shows the procedures used for calculating perfect scores.

Table 5a. Estimation of Perfect Scores

Scores	Estimate (Logits)	Entries	Difference	Perfect Score Value
47	3.22			
46	2.66	0.56		
45	2.33	0.33	0.23	
MAX = 48	3.22	+ 0.56	+ 0.23	= 4.01

The same procedure was employed to calculate zero scores except that the three lowest raw scores and logit values closest to 0 were chosen (i.e. 1, 2 and 3) and subtractions were conducted from the bottom. Table 5b presents the data and the estimated zero score value using this procedure.

Table 5b. Estimation of Zero Scores

Scores	Estimate (Logits)	Entries	Difference	Zero Score Value
3	-4.68			
2	-5.35	-0.67		
1	-6.41	-1.06	-0.39	
MIN 0	-6.41	+ -1.06	+ -0.39	-7.86

The entry -1.06 was estimated by subtracting -5.35 from -6.41, and the entry −0.67 was obtained by subtracting -4.68 from -5.35. The difference -0.39 was estimated by subtracting -0.67 from -1.06, while the zero score value of -7.86 was estimated by adding -6.41 and -1.06 and -0.39.

The above section discusses the procedures for calculating scores of the Chinese achievement and English Word Knowledge tests using the Rasch model. The main purposes of calculating these scores are to: (a) examine the mean levels of all students' achievement in learning the Chinese language between grade levels and across term occasions; (b) provide data on the measures for individual students' achievement in learning the Chinese language between terms for estimating individual student's growth in learning the Chinese language over time; and (c) test the hypothesised models of student-level factors and class-level factors influencing student achievement in learning the Chinese language. The following section considers the procedures for equating the Chinese achievement tests between grades and across terms, as well as the English word knowledge tests across grades.

Equating of the Chinese Achievement Tests between Terms

Table 2 shows the number of anchor items across terms and bridge items between grades as well as the total number and the number of deleted items. The anchor items were required in order to examine the achievement growth of the same group of students over time, while the bridge items were developed so that the achievement growth between grades could be estimated. It should be noted that the number of anchor items was greater in Terms 2 and 3 than in Terms 1 and 4. This was because the anchor items in Term 2 included common items for both Term 1 and Term 3, and the anchor items in Term 3 included common items for both Term 2 and Term 4, whereas Term 1 only provided common items for Term 2, and Term 4 only had common items from Term 3.

The location of the bridge items in a test remained the same as their location in the lower grade level tests for the same term. For example, the Items 28 to 32 were bridge items between Year 5 and Year 6 in the Term 1 tests, and their numbers were the same in the tests at both levels. The raw responses of the bridge items were entered under the same item numbers in the SPSS data file regardless of different grade levels and terms. However, the anchor items were numbered in accordance with the items in particular grade levels and different terms. This is to say that the anchor items in Year 6 for Term 2 were numbered 10 to 14, while in Term 3 test they might be numbered 12 to 16 depending upon the design of Term 3 test.

It can be seen in Table 2 that the number of bridge items varied slightly between grade levels and occasions. In general, the bridge items at one grade level were common to the two adjacent grade levels. For example, there were 10 bridge items in Year 5 for the Term 2 test. Out of the 10 items, five were from the Year 4 test, and the other five were linked to the Year 6 test. Year 4 only had five bridge items each term because it only provided common items for Year 5.

In order to compare students' Chinese language achievement across grade levels and over terms, the anchor item equating method was employed to equate the test data sets of Terms 1, 2, 3 and 4. This is done by initially estimating the item threshold values for the anchor items in the Term 1 tests. These threshold values were then fixed for these anchor items in the Term 2 tests. Thus, the Term 1 and Term 2 data sets were first equated, followed by equating the Terms 2 and 3 data files by fixing the threshold values of their common anchor items. Finally Terms 3 and 4 data were equated. In this method, the anchor items in Term 1 were equated using anchor item equating in order to obtain appropriate thresholds for all items in Term 2 on the scale that had been defined for Term 1. In this way the anchor items in Term 2 were able to be anchored at the thresholds of those corresponding anchor items in Term 1. The same procedures were employed to equate Terms 2 and 3 tests, as well as Terms 3 and 4 tests. In other words, the threshold values of anchor items in the previous term scores were estimated for equating all the items in the subsequent term. It is clear that the tests for Terms 2, 3, and 4 are fixed to the zero point of the Term 1 tests. Zero point is defined to be the average difficulty level of the Term 1 items used in calibration of the Term 1 data set.

Tables 6 to 8 present the anchor item thresholds used in equating procedures between Terms 1, 2, 3 and 4. In Table 6, the first column shows the number of anchor items in the Term 2 data set, the second column displays the number of the corresponding anchor items in the Term 1 data, and the third column presents the threshold value of each anchor item in the Term 1 data file.

It is necessary to note that Level 5 data were not available for Terms 3 and 4 because the students at this level were preparing for Year 12 SACE examinations. As a consequence, the Level 5 data were not included in the data analyses for Term 3 and Term 4.

Equating of English Word Knowledge Tests

With respect to the three English Word Knowledge tests, Item 11 and Item 95 misfitted the Rasch scale in the analyses of Test 1V and Test 3V. However, when the three tests were combined by common items and analysed in one single file, both items fitted the Rasch scale. Consequently, no item was deleted from the calibration analysis.

Concurrent equating was employed to equate the three English Word Knowledge tests, namely Tests 1V, 2V and 3V. Test 1V was administered to students at Years 4 to 6 and Level 1, Test 2V was administered to students at Levels 2 and 3, and Test 3V was completed by Levels 4 and 5 students. In the process of equating, the data from the three tests were combined into a single file so that the analysis was conducted with one data set. It was noted above in this chapter that all the items fitted the Rasch scale when the combined data were calibrated and analyzed. Therefore, no item was deleted from the calibration in the English word knowledge tests for equating purposes. There were 34 common items between the three tests, of which 13 items were common between Tests 1V and 2V, whereas 21 items were common between Tests 2V

and 3V. Furthermore, all the three test data files shared two of the 34 common items. The thresholds of the 34 items obtained during the calibration were used as anchor values for equating the three test data files and for calculating the Rasch scores for each student. Therefore, the 120 items became 86 items after the three tests were combined into one data file.

Table 6. Description of Anchor Item Equating between Terms 1 and 2

Term 2 items		**Term 1 items**	**Thresholds**
Year 4	item 2	3	anchored at -1.96
	item 3	4	anchored at -1.24
	item 4	5	anchored at -2.61
	item 5	6	anchored at -1.50
Year 5	item 16	18	anchored at 0.46
	item 17	2	anchored at -2.25
	item 18	19	anchored at -0.34
	item 19	20	anchored at 1.04
	item 20	6	anchored at -1.50
Year 6	item 29	22	anchored at -0.72
	item 30	23	anchored at 0.20
Level 1	item 46	47	anchored at 0.36
	item 47	48	anchored at -1.59
	item 48	49	anchored at 1.13
	item 49	50	anchored at 0.22
	item 50	51	anchored at 0.34
Level 2	item 95	66	anchored at -3.22
	item 96	70	anchored at -1.03
	item 97	79	anchored at -1.86
	item 98	78	anchored at -0.97
	item 99	76	anchored at -1.15
Level 3	item 175	126	anchored at -1.40
	item 176	127	anchored at 1.37
	item 177	128	anchored at -0.44
	item 178	129	anchored at -0.86
	item 179	130	anchored at 1.83
Level 4	item 240	142	anchored at -1.38
	tem 241	143	anchored at 0.25
	item 243	144	anchored at -0.50
	item 244	145	anchored at -0.50
Level 5	item 302	137	anchored at 0.53
	item 304	139	anchored at 1.80
	item 306	141	anchored at 2.44
Total 33 items			

Note: Probability level = 0.50

In the above sections, the calibration, equating and calculation of scores of both the Chinese language achievement tests and English Word Knowledge tests are discussed. The section that follows presents the comparisons of students achievement in learning the Chinese language across grade levels and over the four school terms, as well as the comparisons of the English Word Knowledge results across grade levels.

Differences in the Scores on the Chinese Language Achievement Tests

The comparisons of students' achievement in learning the Chinese language were examined in the following three ways: (a) comparisons over the four occasions; (b) comparisons between grade levels; and (c) comparisons within grade levels. English Word Knowledge tests results were only compared across grade levels because the tests were administered on only one occasion.

Table 7. Description of Anchor Item Equating between Terms 2 and 3

	Term 3 items	Term 2 items	Thresholds
Year 4	item 2	2	anchored at -2.25
	item 3	3	anchored at -1.96
	item 4	4	anchored at -1.24
	item 5	5	anchored at -2.61
	item 6	6	anchored at -1.50
	item 7	7	anchored at -4.00
	item 8	8	anchored at -3.95
	item 9	9	anchored at -3.58
	item 10	10	anchored at -1.55
	item 11	11	anchored at -0.25
	item 12	12	anchored at -1.91
Year 5	item 18	21	anchored at -0.34
	item 19	22	anchored at 1.04
	item 20	23	anchored at 0.34
	item 21	24	anchored at -0.31
	item 22	25	anchored at 1.11
Year 6	item 38	31	anchored at -1.49
	item 39	34	anchored at 0.64
	item 40	37	anchored at -0.47
	item 41	39	anchored at -0.35
	item 42	40	anchored at -0.30
Level 1	item 48	41	anchored at 0.45
	item 49	42	anchored at -0.26
	item 50	43	anchored at 0.76
	item 51	44	anchored at 0.61
	item 52	45	anchored at 0.69
Level 2	item 107	100	anchored at 0.28
	item 108	101	anchored at 2.43
	item 111	104	anchored at -0.67
Level 3	item 158	187	anchored at 1.14
	item 159	188	anchored at 1.14
Total 31 items			

Notes: a) Probability level = 0.50 b) The items at Levels 4 and 5 misfitted the Rasch model and were therefore deleted.

Comparisons between Occasions

Table 9 shows the scores achieved by students on the four term occasions, and Figure 1 shows the achievement level by occasions graphically. It is interesting to notice that the figures indicate general growth in student achievement mean score between Terms 1 and 2 (by 0.53), Terms 2 and 3 (by 0.84), whereas an obvious drop in the achievement mean score is seen between Terms 3 and 4 (by 0.17). The drop of achievement level in Term 4 might result from the fact that some students had decided to drop out from learning the Chinese language in the next year, thus they ceased to put effort into the learning of the Chinese language.

Comparisons between Grades on Four Occasions

This comparison was made between grade levels on the four different occasions. After scoring, the mean score for each grade was calculated for each occasion. Table 10 presents the mean scores for the students at Year 4 to Year 6, and Level 1 to Level 5, and shows increased achievement levels between the first three terms. However, the achievement level decreases in Term 4 for Year 4, Year 5, Level 1, Level 3, and Level 4. The highest level of achievement for these grades is, in general, on the Term 3 tests. The achievement level for students in Year 6 is higher for Term 1 than for Term 2. However, sound growth is observed between Term 2 and Term 3, and Term 3 and Term 4.

Table 8. Description of Anchor Item Equating between Terms 3 and 4

	Term 4 items	Term 3 items	Thresholds
Year 4	item 2	3	anchored at -1.96
	item 3	4	anchored at -1.24
	item 4	5	anchored at -2.61
	item 5	6	anchored at -1.50
Year 5	item 26	28	anchored at 0.70
	item 27	29	anchored at 1.12
	item 28	30	anchored at 0.25
	item 29	31	anchored at 1.27
	item 30	32	anchored at 1.17
Year 6	item 31	30	anchored at 0.25
	item 32	31	anchored at 1.27
	item 33	32	anchored at 1.17
	item 34	43	anchored at -0.48
	item 35	44	anchored at 1.30
Level 1	item 41	58	anchored at 2.03
	item 42	59	anchored at 2.60
	item 43	60	anchored at 1.79
	item 44	61	anchored at 2.37
	item 45	62	anchored at 1.60
	item 46	63	anchored at 0.38
	item 4	64	anchored at 0.84
	item 48	65	anchored at 0.80
	item 49	66	anchored at 0.48
	item 50	67	anchored at 0.59
Level 3	item 185	165	anchored at 5.46
Level 4	item 234	188	anchored at 2.38
	item 236	190	anchored at 3.41
Total 27 items			

Notes: a) Probability level = 0.50; b) The items at Level 2 misfitted the Rasch model and were therefore deleted. Level 5 tests were not available for Terms 3 and 4.

It is of interest to note that the students at Level 2 achieved a marked growth between Term 1 and Term 2, namely from -0.07 to 2.27. The highest achievement level for this grade is at Term 4 with a mean score of 2.88. Students at Level 4 are observed to have achieved their highest level at Term 3.

Table 9. Average Rasch Scores on Chinese Achievement Tests by Term

Term 1	Term 2	Term 3	Term 4
0.43	0.96	1.80	1.63
N=781	N=804	N=762	N=762

The lowest achievement level for this grade is at Term 2. Because of the inadequate information provided for Level 5 group, it was not considered possible to summarise the achievement level for that grade. Figure 2 presents the differences in the achievement levels between grade levels on four occasions based on the scores for each grade as well as for each term in Table 10. Figure 3 below presents the mean differences in the achievement levels between grades over the four occasions.

Figures 1, 2 and 3 present graphically the achievement levels for each grade for the four terms. Figure 1 provides a picture of students' achievement level on different occasions, while Figures 2 and 4 show that there is a marked variability in the achievement level across terms between and within grades. However, the general trend of a positive slope is seen for Term 1 in Figure 2. A positive slope is also seen for performance at Term 2 despite the noticeable drop at Level 4. The slope of the graph for Term 3 can be best described as erratic because a large decline occurs at Year 6 and a slight decrease occurs at Level 2. It is important to note that the

trend line for Term 4 reveals a considerable growth in the achievement level although it declines markedly at Level 4.

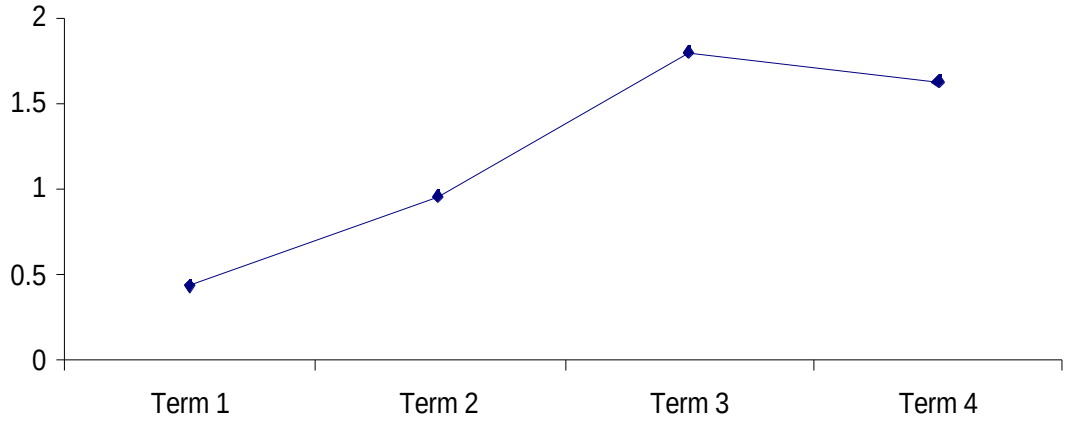


Figure 1. Chinese Achievement Level by Four Term Occasions

Table 10. Average Rasch Scores on Chinese Tests by Term and by Year Level

LEVEL	Term 1	Term 2	Term 3	Term 4	Mean
Year 4	-0.93	-0.15	0.35	0.09	-0.17
Year 5	0.20	0.52	1.86	1.47	1.01
Year 6	0.81	0.69	0.85	1.06	0.85
Level 1	0.73	0.95	2.33	1.77	1.45
Level 2	-0.07	2.27	2.13	2.88	1.80
Level 3	1.71	2.65	4.62	4.30	3.32
Level 4	1.35	0.31	4.64	1.86	2.04
Level 5	2.43	1.58	-	-	-
No. of cases	N=782	N=804	N N=762	N=762	-

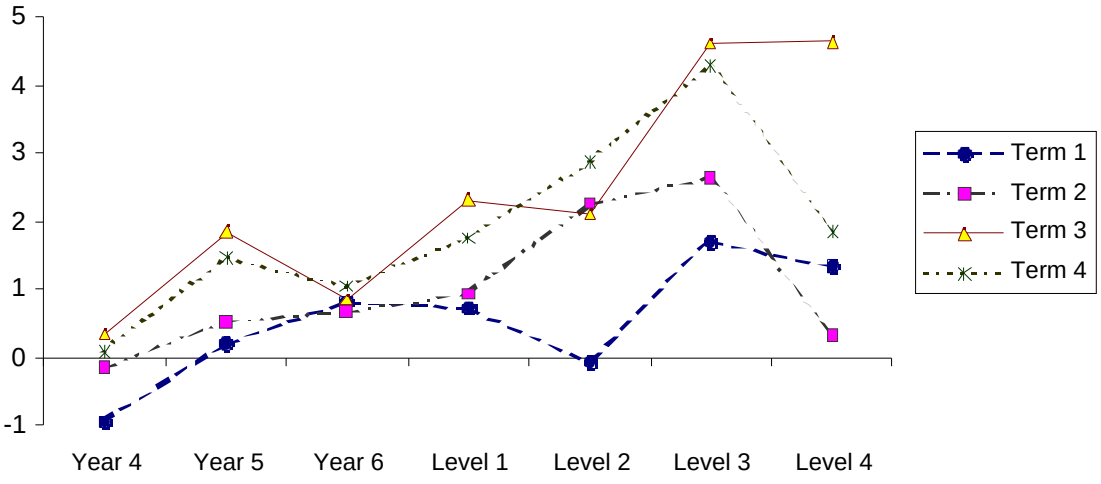


Figure 2. Description of Achievement Level between Grades on Four Occasions

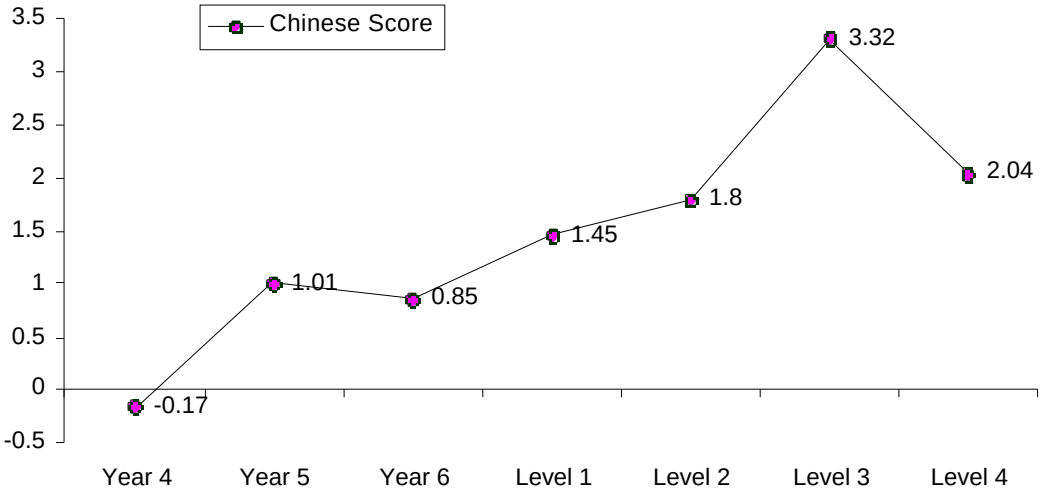


Figure 3. Comparison of Means in Chinese Achievement between Grade Levels

Figure 3 presents the comparisons of the differences in student achievement level between grades. It is of importance to note that students at Level 3 achieved the highest level among the seven grade levels, followed by Level 4, while students at Year 4 were the lowest as might be expected.

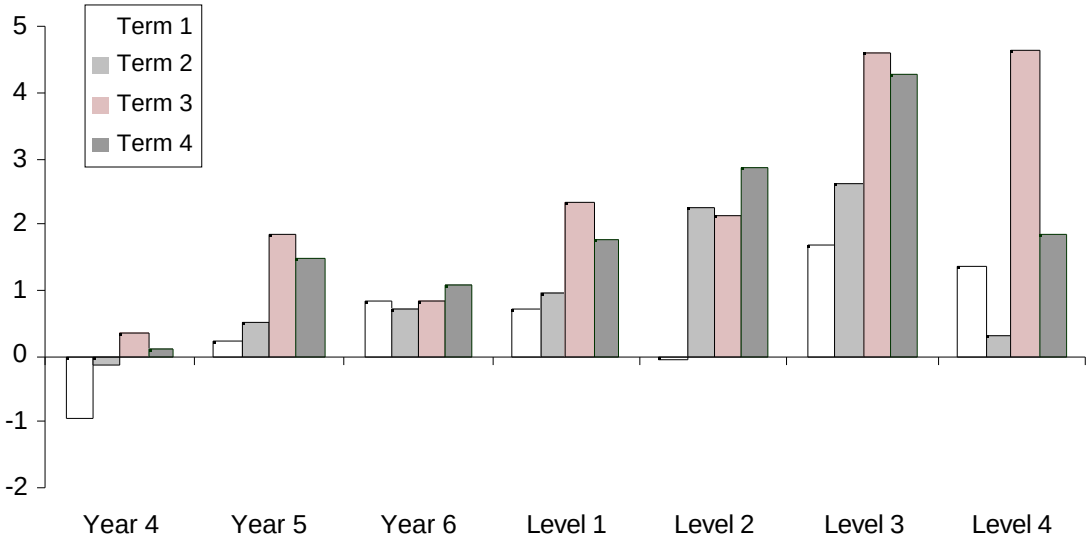


Figure 4. Description of Achievement Level within Grades on Four Occasions

This might be explained by the fact that four of the six Year 4 classes learned the Chinese language only for two terms, namely, Terms 1 and 2 in the 1999 school year. They learned French in Terms 3 and 4. Figure 4 shows the achievement levels within grade levels on the four occasions, which are discussed in the following section.

Comparisons within Grades on Different Occasions

This section compares the achievement level within each grade. By and large, an increased trend is observed for each grade level from Term 1 to Term 4 (see Figures 1, 2 and 4, and Table 10). Year 4 students achieve at a markedly higher level between Terms 1, 2 and 3. The increase is 0.78 between Term 1 and Term 2, and 0.20 between Term 2 and Term 3. However, the decline between Term 3 and Term 4 is 0.26. Year 5 is observed to show a similar trend in the achievement level as Year 4. The growth difference is 0.32 between Term 1 and Term 2. A

highly dramatic growth difference is seen of 1.34 between Term 2 and Term 3. Although a decline of 0.39 is observed between Term 3 and Term 4, the achievement level in Term 4 is still considered high in comparison with Terms 1 and 2.

The tables and graphs above show a consistent growth in achievement level for Year 6 except for a slight drop at Term 2. The figures for achievement at Level 1 reveal striking progress at Term 3 followed by Term 4, and a consistent growth is shown between Terms 1 and 2. At Level 2 while a poor level of achievement is indicated at Term 1, considerably higher levels are achieved for the subsequent terms. The students at Level 3 achieve a remarkable level of performance across all terms even though a slight decline is observed at Term 4. The achievement level at Level 4 appears unstable because a markedly low level and extremely high level are achieved at Term 2 and Term 3, respectively.

Despite the variability in the achievement levels on the different occasions at each grade level, a common trend is revealed with a decline in achievement that occurs at Term 4. This decline might have resulted from the fact that some students had decided to drop out from the learning of the Chinese language for the next year, and therefore ceased to put in effort in Term 4. With respect to the differences in the achievement level between and within grades on different occasions, further analyses are required.

Difference in English Word Knowledge Tests between Grade Levels

Three English Word Knowledge tests were administered to students who participated in this study in order to investigate the relationship between the Chinese language achievement level and proficiency in the English Word Knowledge tests. Table 11 provides the results of the scores by each grade level, and Figure 5 presents graphically the trend in English Word Knowledge proficiency between grade levels using the scores recorded in Table 11.

Table 11. Average Rasch Score on English Word Knowledge Tests by Grade

Level	No. of Students (N)	Scores
Year 4	154	-0.20
Year 5	167	0.39
Year 6	168	0.63
Level 1	158	0.70
Level 2	105	1.13
Level 3	46	1.33
Level 4	22	1.36
Level 5	22	2.07
Total	842 (103 cases missing)	Mean = 0.93

Table 11 presents the mean Rasch scores on the combined English word knowledge tests for the eight grade levels. It is of interest to note the general improvement in English word knowledge proficiency between grades. The difference is 0.59 between Years 4 and 5; 0.24 between Years 5 and 6; 0.07 between Year 6 and Level 1; 0.43 between Levels 1 and 2; 0.20 between Levels 2 and 3; a small difference of 0.03 between Levels 3 and 4; and a large increase between Levels 4 and 5.

It is also of interest to note the marked development in the English Word Knowledge proficiency between grade levels. Large differences occur between Years 4 and 5, as well as between Levels 4 and 5. Medium or slight differences occur between other grades, namely between Years 5 and 6; Year 6 and Level 1; Levels 1 and 2; Levels 2 and 3; and Levels 3 and 4. The differences between grade levels, whether large or small, are to be expected because as students grow older and move up a grade, they learn more words and develop their English vocabulary and thus may be considered to advance in verbal ability.

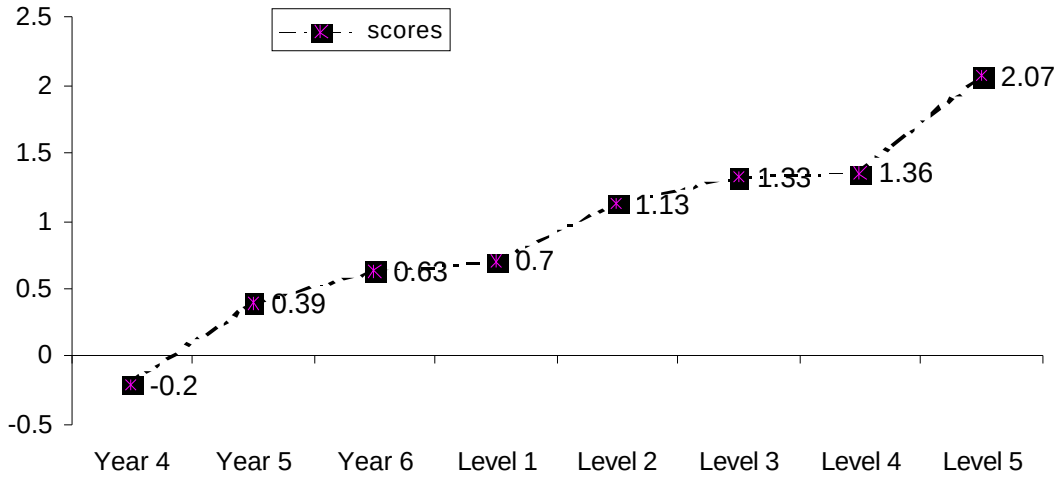


Figure 5. Graph of Scores on English Word Knowledge Tests across Grade Levels

In order to examine whether the development in the Chinese language is associated with the development in English Word Knowledge proficiency as well as to investigate the interrelationship between the two languages, students' achievement level in the Chinese language and development of English Word Knowledge (see Figures 3 and 5) are combined to produce Figure 6. The combined lines indicate that the development of the two languages is by and large interrelated except for the drops at Year 6 from Year 5 and at Level 4 from Level 3 in the level of achievement in learning the Chinese language. This suggests that both students' achievement in the Chinese language and development in English Word Knowledge proficiency generally increase across grade levels.

It should be noted that both sets of scores are recorded on logit scales in which the average levels of difficulty of the items within the scales determine the zero or fixed point of the scales. It is thus fortuitous that the scale scores for the students at Year 4 for English word knowledge proficiency and performance on the tests of achievement in learning the Chinese language are so similar.

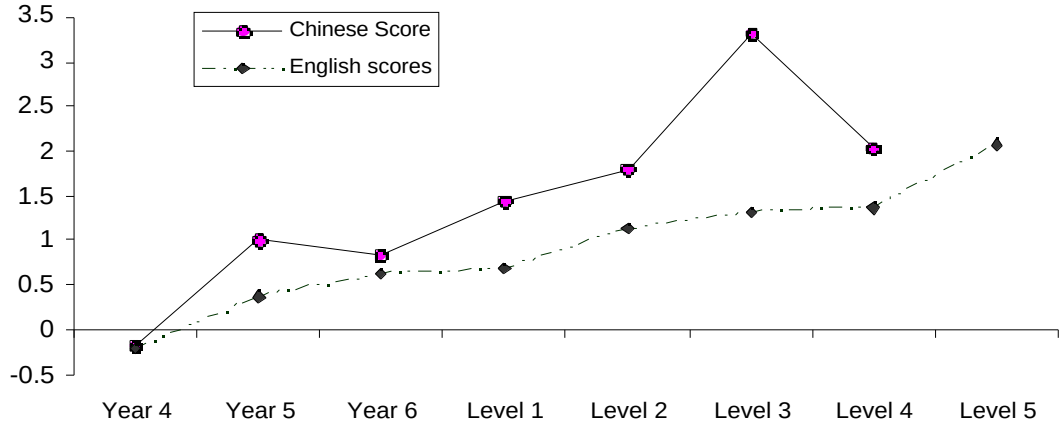


Figure 6. Comparison between Chinese and English Scores by Grade Level

CONCLUSION

The findings from the Chinese achievement tests and English Word Knowledge tests can be summarised as follows. First, the students' achievement level in the Chinese language generally improves between occasions though there is a slight decline at Term 4. Second, overall, the higher the grade level in the Chinese language, the higher the achievement level.

Third, within-grade achievement level in learning the Chinese language for each grade level indicates a consistent improvement across the first three terms but shows a decline at Term 4. The decline in performance for Term 4, particularly of students at Level 4, may have been a consequence of the misfitting essay type items that were included in the tests at this level, which resulted in an underestimate of the students' scores. Finally, the achievement level in learning the Chinese language appears to be associated with the development of English word knowledge, namely, the students at higher grade levels are higher achievers in both English word knowledge and on the Chinese language tests.

Although a common trend is observed in the findings for both the Chinese language achievement and English word knowledge development, differences still exist within and between grade levels as well as across terms. It is considered necessary to investigate what factors gave rise to such differences.

Nevertheless, the work of calibrating and equating so many tests across so many different levels and so many occasions is to some extent a hazardous task. There is the clear possibility of errors in equating, particularly in situations where relatively few anchor items and bridge items are being employed. However, stability and strength are provided by the requirement that items must fit the Rasch model, which in general they do well.

REFERENCES

Adams, R. & Khoo, S-T. (1993). *Quest: The Interactive Test Analysis System*. Melbourne: ACER.

Afrassa, T.M. (1998). *Mathematics achievement at the lower secondary school stage in Australia and Ethiopia: A comparative study of standards of achievement and student level factors influencing achievement.* Unpublished Doctoral Thesis. School of Education, The Flinders University of South Australia, Adelaide.

Andrich, D. (1988). *Rasch Models for Measurement*. Series: Quantitative applications in the social sciences. Newbury Park, CA: Sage Publications.

Andrich, D. & Masters, G.N. (1985). Rating scale analysis. In T. Husén and T. N. Postlethwaite (eds.), *The International Encyclopedia of Education.* Oxford: Pergamon Press, pp. 418-4187.

Angoff, W. H. (1982). Summary and derivation of equating methods used at ETS. In P. W. Holland and D. B. Rubin (eds.), *Test Equating*. New York: Academic Press, pp. 55-69.

Auchmuty, J.J. (Chairman) (1970*). Teaching of Asian Languages and Cultures in Australia.* Report to the Minister for Education. Canberra: Australian Government Publishing Service (AGPS).

Baker, F.B. and Al-Karni (1991). A comparison of two procedures for computing IRT equating coefficients. *Journal of Educational Measurement,* 28 (2), 147-162.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. Lord and M.Novick, *Statistical Theories of Mental Test Scores.* Reading MA: Addison-Wesley, pp.397-472.

Carroll, J.B. (1975). *The Teaching of French as a Foreign Language in Eight Countries.* International Association for the Evaluation of Educational Achievement (IEA). Stockholm, Sweden: Almqvist & Wiksell.

Commonwealth Advisory Committee (1970). *The Teaching of Asian Languages and Cultures in Australia (The Auchmuty Report).* Commonwealth Advisory Committee. Canberra: AGPS.

Commonwealth of Australia (1998). *Advancing Australia's Languages*: *Overview Report*. Canberra: The Evaluation Team, Faculty of Asian Studies, The Australian University.

Everson, M. E. & Ke, C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *JCLTA,* 32 (1), 1-20.

Fairbank, K. & Pegalo, C. (1983). *Foreign Languages in Secondary Schools*. Queensland: Queensland Department of Education.

Holland, P.W. & Rubin, D.B. (1982). *Test Equating*. New York: Academic Press.

Hyltenstam, K. & Pienemann, M. (eds.) (1985). *Modeling and Assessing Second Language Acquisition.* Clevedon: Multilingual Matters.

Keeves, J.P. (1992b). Scaling achievement test scores. In J.P. Keeves (ed.), *The IEA Technical Handbook.* The Hague, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA), pp. 107-125.

Keeves, J.P. (1997). Models and model building. In J. P. Keeves (ed.): *Educational Research, Methodology, and Measurement: An International Handbook,* (*2nd edn.*). Oxford: Pergamon Press, pp.386-394.

Keeves, J.P. & Alagumalai, S. (1999). New approaches to measurement. In G. N. Masters and J. P Keeves (eds.) *Advances in Measurement in Educational and Psychological Research and Assessment.* Oxford: Pergamon, pp. 23-42.

Kenyon, D.M. & Stansfield, C.W. (1992). Extending a scale of language proficiency using concurrent calibration and the Rasch Model. *ERIC, ED 343- 443,* 1-26.

Kenyon, D.M. & Stansfield, C.W. (1992). Examining the validity of a scale used in a performance assessment from many angles using the Many-Faceted Rasch Model. *ERIC, ED 343 442,* 1-39.

Kolen, M.J.(1997). Equating of tests. In J.P. Keeves (ed.), *Educational Research, Methodology, and Measurement: An International Handbook (2nd edn.)*. Oxford: Pergamon Press, pp.730-737.

Lietz, P. (1995). *Changes in reading comprehension across cultures and over time*. Unpublished Doctoral Thesis. School of Education, The Flinders University of South Australia, Adelaide.

Linacre, J.M. (1989). *Many-facet Rasch Measurement*. Chicago: MESA Press.

Masters, G.N. (1997). Partial credit model. In J. P. Keeves (ed.): *Educational Research, Methodology, and Measurement: An International Handbook,* (*2nd edn.*). Oxford: Pergamon Press, pp.292-299.

Masters, G.N. & Keeves, J.P. (1999). *Advances in Measurement in Educational and Psychological Research and Assessment.* Oxford: Pergamon Press.

McNamara, T.F. (1996). *Measuring Second Language Performance*. New York: Addison Wesley Longman.

McNamara, T.F. (1997). 'Interaction' in second language performance assessment: Whose performance? *Applied Linguistics,* 18 (4), 446-466.

Mohandas, R. (1996). *Test equating: problems and solutions*. Unpublished M.Ed thesis. School of Education, The Flinders University of South Australia, Adelaide.

Morrison, C.A. & Fitzpatrick, S.J. (1992). *Direct and Indirect Equating*: *A Comparison of Four Methods Using the Rasch Model.* Austin, Texas: Measurement and Evaluation Center, the University of Texas.

Noru_is, M.J. (1993). SPSS for Windows: Advanced Statistic Release 6.0. Chicago, IL: SPSS Inc.

Murray, D. & Lundberg, K. (1976). *A Register of Modern Language Teaching in South Australia*. INTERIM REPORT, Document No. 50/76, Adelaide.

Petersen, N.S.; Marco, G.L.; & Stewart, E E. (1982). A test of the adequacy of linear score equating models. In P.W. Holland and D.B. Rubin (eds.), *Test Equating*. New York: Academic Press, pp. 71-135.

Scarino, A.; Vale, D.; McKay, P. & Clark, J. (1988). *Australian Languages Levels Guidelines: Book 1, Language Learning in Australia.* Canberra: Commonwealth of Australia.

Smith, D., Chin, N.B., Louie, K., & Mackerras, C. (1993). *Unlocking Australia's Language Potential: Profiles of 9 Key Languages in Australia, Vol. 2: Chinese.* Canberra: Commonwealth of Australia and NLLIA.

Stocking, M.L. (1997). Item response theory. In J. P. Keeves (ed.), *Educational Research, Methodology, and Measurement: An International Handbook,* (*2nd edn.*), Oxford: Pergamon Press, pp. 836-840.

Thorndike, R.L. (1973b). Reading as reasoning. *Reading Research Quarterly,* 9 (2), 135-147.

Tuffin, P. & Wilson, J. (1990). *Report of an Investigation into Disincentives to Language Learning at the Senior Secondary Level.* Commissioned by the Asian Studies Council, Adelaide.

Weiss, D.J. & Yoes, M.E. (1991). Item response theory. In R. K. Hambleton and J. N. Zaal (eds.), *Advances in Education and Psychological Testing.* Boston, MA: Kluwer Academic Publishers, pp. 69-96.

Wright, B.D. (1985). Rasch measurement models. In T. Husén and T N. Postlethwaite (eds.), *The International Encyclopedia of Education*. Oxford: Pergamon Press, pp. 4177-4181.

Wright, B.D. (1988). *George Rasch and Measurement*. Informal remarks at the Inaugural Meeting of the American Educational Research Association Rasch Measurement SIG, New Orleans, April 8.