International Education Journal

International Education Journal

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.

Copyright - Copyright of material produced in this journal rests with individual contributors. Apart from fair use such as brief quotation for scholarly purposes, permission to use material in this journal should be obtained from the authors concerned.

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 to submit 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.

Author Information

IEJ welcomes practical and research manuscripts that focus on educational issues, provide a clear purpose and depth of discussion, and are presented 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, 13 Endeavour Court, North Haven, SA 5018.

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).

  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.

Differential item functioning: A case study from first year economics

Steven Barrett
University of South Australia *steven.barrett@unisa.edu.au*

The Division of Business and Enterprise at the University of South Australia is attempting to reduce teaching costs. Consequently, multiple-choice tests are set to play a more important role in assessment processes. However, many subject coordinators do not have the information required to ensure that this form of assessment is both effective and efficient.

The study aims to determine whether a recent multiple-choice examination effectively discriminated between students on the basis of ability or were gender biased. The extent of guessing was also investigated. The study is an application of Item Response Theory using the Rasch Model. The study found little evidence of gender bias or systematic guessing, but did determine that many questions did not adequately discriminate between students of different ability.

Item Response Theory, Rasch model

INTRODUCTION

Like many other Australian business programs, the Division of Business and Enterprise at the University of South Australia is teaching more students with increasingly fewer resources. One cost saving initiative is the introduction of the 'faculty core'. All commencing undergraduate business students study eight common first year subjects. These subjects are taught in large lecture theatres, with fortnightly workshops of 30 to 36 students, unstaffed study groups in the 'off weeks' and extensive use of casual staff. It is therefore not surprising that subject conveners are increasingly using multiple-choice examinations, as the cost savings are significant (Walstead and Becker 1994).

Multiple-choice questions are primarily obtained from the instructor's manuals and test banks that accompany business textbooks. If these items are to form the basis of examinations that reflect accurately the learning outcomes of students, then a range of questions need to asked about them. The first question is obviously, who writes these questions and what is their educational background? The answer, at least in first year economics subjects, tends to be, that American graduate students, who are paid on a per question basis, write them. Anecdotal evidence gleaned from textbook company representatives suggest that about US$2.00 per question is the going rate. This observation in turn gives rise to a raft of other questions about their validity as testing instruments. For example, do they adequately discriminate between students on the basis of ability, are they biased on the basis of gender or other student characteristics and do students rely on guessing to answer them? The literature supports the view that multiple-choice tests in economics are biased towards males. Walstead and Robson (1997) argue that female students of equal ability do not perform as well as males on multiple-choice tests. Their study could not attribute this discrepancy to sexist language or stereotypical examples. Rather, they suggested that female high school students experience more difficulties with questions involving numerical,

spatial or higher reasoning skills. There is also often a nagging doubt in the minds of lecturers that they do not reflect the true ability of students as a result of guessing (Keats and Lord 1962). Unfortunately, answers to these questions are not readily found in the literature. If questions sourced from instructor's manuals and test banks are biased, do not discriminate between students or are easily guessed correctly, then they are not suitable as a basis for assessment.

This paper analyses the 30 multiple-choice questions that were set as part of the final examination for Economic Environment in Semester 1, 1997. The study attempted to find answers to three questions. First, do these questions discriminate between students on the basis of their ability? The second question reflects a growing body of literature that argues that economics courses are not gender inclusive. Such gender blindness may be carried over into the assessment process. Hence, it was decided to determine if these are the questions biased in favour of males? The third question is based on the observation that business subjects make extensive use of multiple-choice questions. Yet the terms multiple-choice and multiple-guess are used synonymously. Hence, it was decided to investigate whether students consistently guessed the answers to the multiple-choice questions on this paper.

METHODS

This study is an analysis of the performance of the 844 students who attempted the multiple-choice section of the final examination for Economic Environment in Semester 1, 1997. The paper comprised 30 multiple-choice questions, a short answer section and an essay section, each of which accounted for one third of the marks on the paper. The performance of these students was analysed using the Rasch Model. The item analyses were obtained by using the software package Quest (Adams and Khoo 1993) and the Item Characteristic Curves were obtained from RUMM (Rasch Unidimensional Measurement Models (Sheridan, Andrich and Luo 1997).

Quest and RUMM. The mathematical models that underlie these two packages differ slightly and hence the results of the two packages also differ slightly. The following discussion is based largely on the output obtained from Quest. However, the output from RUMM has the advantage that it can be graphed. In particular the Item Characteristic Curves can be obtained for each question. These curves provide a much clearer picture of the way that questions discriminate between students of different ability.

THEORETICAL FRAMEWORK

The results of the multiple-choice examination were analysed using the Rasch Model, which is a latent trait model based on the premise that the performance of students is determined by an underlying, or latent trait, that is not observable. The latent trait that enables students to answer correctly a question on a multiple-choice test is usually referred to as ability. The model is probabilistic and is concerned with defining and predicting the probability of obtaining a correct answer to a test question as a function of the underlying trait. The model attempts to develop and specify the relationship between the observable performance of students in an examination and the latent trait.

The Rasch Model predicts the odds, or probability, of a student obtaining a correct answer on a multiple-choice test in terms of two parameters, one relating to the difficulty of the items on the test and the other to the ability of the students. The basis of the model is that the relationship between item difficulty and student ability determines the performance of students on a test. That is, a student with greater ability should also have a higher chance of success on a particular question than a less able person. Conversely, a person of any level of ability would have a greater chance of success on a less difficult question than on a more difficult question. To the extent that

this relationship holds, the probability of the success of a student on a question can be specified as a function of the difference between the ability of the student and the difficulty of a question, when both ability and difficulty are measured on the same linear scale using the same units of measurement. Item Response Theory and the Rasch Model facilitates the construction of an interval scale that allows two disparate concepts such as student ability and item difficulty to be measured and compared.

The relationship between ability and difficulty can be expressed simply in the forms of odds. The odds of getting a question correct (O) is equal to the ability of the student (A) multiplied by the easiness of the question (E), that is, O = AE. To illustrate this relationship, if ability was zero then all questions would be impossible. On the other hand, if ability were very high, then all but the most difficult questions would be easy. At the mid-point between these extremes of student ability and easiness of the question are reciprocals such that AE = 1. As this is a probabilistic model the range of both ability and easiness is from zero to one. Hence, the student and the question are perfectly matched and the odds of success would be 1:1 and the probability of success would be 0.5.

If logs of both sides of the relationship are taken, then an additive form of the equation is obtained, logObserved = logAbility + logEasiness. The difficulty of the question can now be substituted for easiness. The use of a difficulty rather than an easiness parameter means that the odds of a correct answer falls as the difficulty of the question rises. This gives, logObserved = logAbility - logDifficulty. The Rasch Model can now be expressed in probabilistic terms;

$$\text{Probability}_{(\text{correct response})} = \frac{\exp(A-D)}{1 + \exp(A-D)}$$

This is the expression for the probability of a single student getting a particular question correct. The expression is then summed across all questions and all students in order to develop the model. Finally the performance of students can be explained in terms of only two parameters, one for the student (ability) and one for the item (difficulty). More importantly, both parameters are expressed in the same cardinal units (logits) and measured on the same scale.

Use of the Rasch Model involves the collection of the responses to a set of test items, such as a multiple-choice exam and an estimation of the values in the parameters in the model for the items and the students that best fit the data. In this study the one parameter Rasch Model was employed. Iterative computer procedures are used to calculate the maximum likelihood estimates of the parameters. Initial estimates are made for item difficulties based on the number of correct answers. Then initial estimates are made for the abilities of the students based on their scores. The initial estimates of ability are then used to improve the estimates of item difficulty, which in turn are used to improve the estimates of student ability. The process is iterated to maximise the fit of the parameter estimates to the test data.

RESULTS

The Rasch Model allows the relationship between item difficulty and student ability to be investigated by employing a mathematical model that converts both difficulty and ability to the same units (logits). Ability and difficulty can then be plotted on a graph using the same interval scale. The Rasch Model sets zero at the mean of the item difficulty. Figure 1 is a plot of the item estimates for the 844 students and 30 items on the test. The left hand side of the figure shows that student ability is roughly normally distributed around the logit value of zero, with a maximum of three and a minimum of minus two and half. The difficulty of the items is also distributed around the mean of zero, by definition, but is bunched in the range of positive one to minus one. There are

only two relatively easy questions and three relatively hard questions. Hence, the test did not discriminate between students at the extremes of the ability scale.

```
Logit Scale
   3.0                         XXX   |
                                     |
                                     |
                               XXX   | 15
                                     |
                           XXXXXXXX  |
   2.0                               | 7
                             XXXXX   |
                                     |
                        XXXXXXXXXXXX  |
                        XXXXXXXXXXXX  |
                        XXXXXXXXXXXX  | 2
   1.0                  XXXXXXXXXXXX  |
                        XXXXXXXXXXXX  | 9 14 23
                       XXXXXXXXXXXXXX | 11 12 22
                     XXXXXXXXXXXXXXXX |
                       XXXXXXXXXXXXXX | 6
                        XXXXXXXXXXXX  | 28
   0.0               XXXXXXXXXXXXXXXX | 10 13 17 19 26
                        XXXXXXXXXXXX  | 30
                       XXXXXXXXXXXXXX | 25
                  XXXXXXXXXXXXXXXXXXX | 5 29 8 24 27
                          XXXXXXXXXX  | 20 21
                        XXXXXXXXXXXX  | 4
  -1.0                   XXXXXXXXXX   |
                          XXXXXXXXXX  | 3 16
                             XXXXX    |
                             XXXX     | 18
                                      | 1
                              XXX     |
  -2.0                                |
                               XX     |
                                      |
                               X      |
                                      |
                                      |
  -3.0                                |
```

Note: Each X represents 3 students.

Figure 1: Item Estimates

DISCRIMINATION

A good test should consistently discriminate between higher and lower ability students. This is a necessary condition if the test as a whole is able to discriminate between students of different abilities. The item analysis feature of Quest provides an overview of the way in which every item was answered, whether the item fits the Rasch Model and the extent to which the item discriminates between students. A value of the discrimination coefficient of 0.20 is considered to be the threshold for acceptance. Items with a value lower than this should modified or deleted from the test. The values of the discrimination coefficient obtained in this study varied from 0.10 to 0.56. Clearly some questions discriminate well on the basis of student ability, while others leave something to be desired. Examples of the item analysis obtained from Quest for two questions are presented in Tables 1 and 2. The corresponding Item Characteristic Curves are shown in Figures 2 and 3.

Questions 1 and 12 are examples of good items that should remain in the examination as they both have a discrimination coefficient of 0.40. However, it is not clear from this single statistic the processes that students employed to answer these two questions. Herein lies the value of RUMM. RUMM can be used to plot the Item Characteristic Curves for each question, which provides a visual impression of the process that students employ to answer each question. Figure 2 shows that weaker students were able to guess the correct answer to Question 1, as indicated by the large positive vertical intercept, the pseudo-guessing parameter, of about 0.250. The question discriminates between students in the lower ability range. However, the question does not discriminate between students in the upper range of ability, as shown by the location of the steepest part of the curve occurring in a narrow range of the negative section of the ability range. On the other, hand very few weak students correctly guessed the answer to Question 12 (Figure 3), as shown by the vertical intercept being very close to zero. The question did not discriminate well between students in the lower ability range. But the item did discriminate quite well in the middle ability range. Hence, students approached two questions with identical discrimination coefficients very differently.

Table 1: Item Analysis for Question 1

Item 1:		Infit MNSQ = 0.93			
		Disc = 0.40			
Categories	1	2	3*	4	5
Count	20	15	709	73	25
Percent (%)	2.4	1.8	84.2	8.7	3.0
Pt-Biserial	-0.05	-0.19	0.40	-0.28	-0.20
p-value	.056	.000	.000	.000	.000
Mean Ability	-0.14	-1.21	0.40	-0.69	-0.89
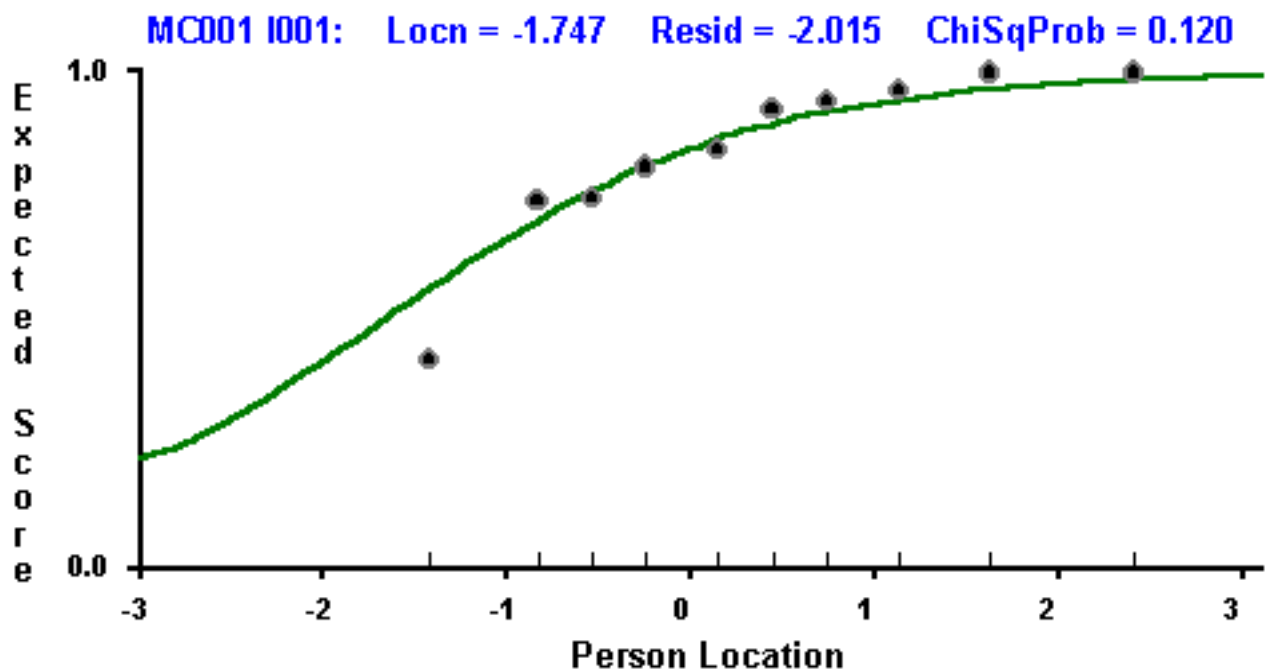
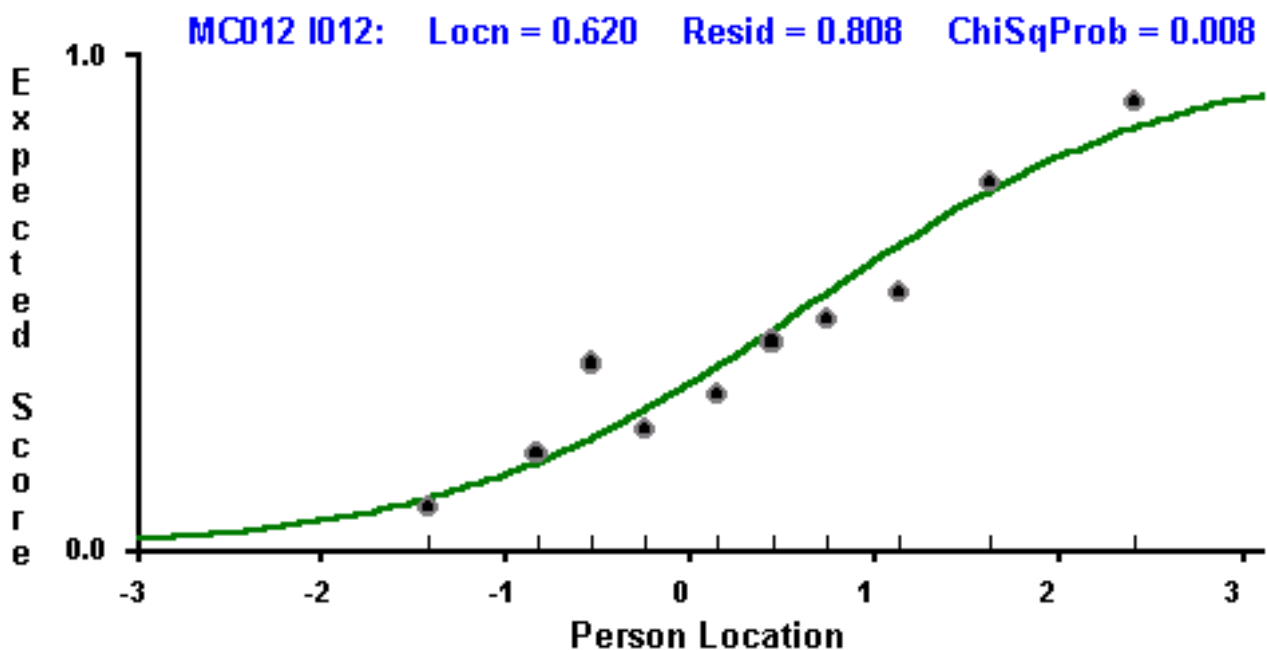**Figure 2: Item Characteristic Curve for Question 1**

RUMM rated this test as good, on a five-point scale from excellent to unacceptable. Quest demonstrated that all questions fitted the Rasch Model, as the Infit Mean Square of every question does not vary significantly from 1.00. Figure 4 shows that the Infit Mean Square for every item. Note that no item lies outside of the threshold range of 0.77 to 1.30.

Table 2: Item Analysis for Question 12

Item 12:		Infit MNSQ = 1.03			
		Disc = 0.40			
Categories	1	2*	3	4	5
Count	72	349	88	118	216
Percent (%)	8.5	41.4	10.4	14.0	25.6
Pt-Biserial	-0.21	0.40	-0.18	0.05	-0.23
p-value	.000	.000	.000	.076	.000
Mean Ability	-0.46	0.72	-0.29	0.33	-0.19

Figure 3: Item Characteristic curve for Question 12

A discrimination coefficient of 0.20 is considered to the threshold and questions below this figure should be deleted. In all but one question (Item 15) the discrimination coefficient exceeded 0.20. On this basis alone the question should be dropped. However, a closer examination of the question reveals that it cannot be answered correctly solely on the basis of material presented in the textbook or discussed in lectures. Hence, this question should be dropped on educational grounds as well. The test indicates that there is something of a gray area in the range between 0.20 and 0.35. Some questions in this range discriminate well while other do not. This suggests that items and their distracters should be closely examined in order to improve the effectiveness of these questions.

Gender Bias

Quest can also be used to determine whether bias exists between identifiable groups of students. Universities collect and record information about a wide range of student characteristics, such as place of birth, aboriginality, location of normal residence, language usually spoken at home, migration history, disability status, employment status and enrolment status. All of which can be used to identify two mutually exclusive and exhaustive groups of students whose performance can be tested for bias. However, this study looked only at bias between male and female students due to the unavailability of data. Nevertheless, the process used to identify bias regardless of its cause is identical. Quest produces a plot of standardised differences between the performance of the two

groups of interest (Figure 5). Items that have a value greater than plus or minus two indicate significant differences in performance between the two groups and the relevant questions need to be examined in order to identify the cause of the bias.

```
              0.63        0.71        0.83        1.00        1.20        1.40        1.60
         --------+---------+---------+---------+---------+---------+---------+---------+-
         item 1                  .                    *   |                     .
         item 2                  .                        |        *            .
         item 3                  .                        |    *                .
         item 4                  .                        |*                    .
         item 5                  .                    *   |                     .
         item 6                  .                        |    *                .
         item 7                  .                        |    *                .
         item 8                  .                     *  |                     .
         item 9                  .                        |*                    .
         item 10                 .                        | *                   .
         item 11                 .                     *  |                     .
         item 12                 .                        |*                    .
         item 13                 .                    *   |                     .
         item 14                 .                        |    *                .
         item 15                 .                        |       *             .
         item 16                 .          *             |                     .
         item 17                 .                      * |                     .
         item 18                 .                     *  |                     .
         item 19                 .                 *      |                     .
         item 20                 .                   *    |                     .
         item 21                 .                      * |                     .
         item 22                 .                        |   *                 .
         item 23                 .                        |*                    .
         item 24                 .                   *    |                     .
         item 25                 .                        |*                    .
         item 26                 .                        |      *              .
         item 27                 .             *          |                     .
         item 28                 .          *             |                     .
         item 29                 .                        |*                    .
         item 30                 .                      * |                     .
```
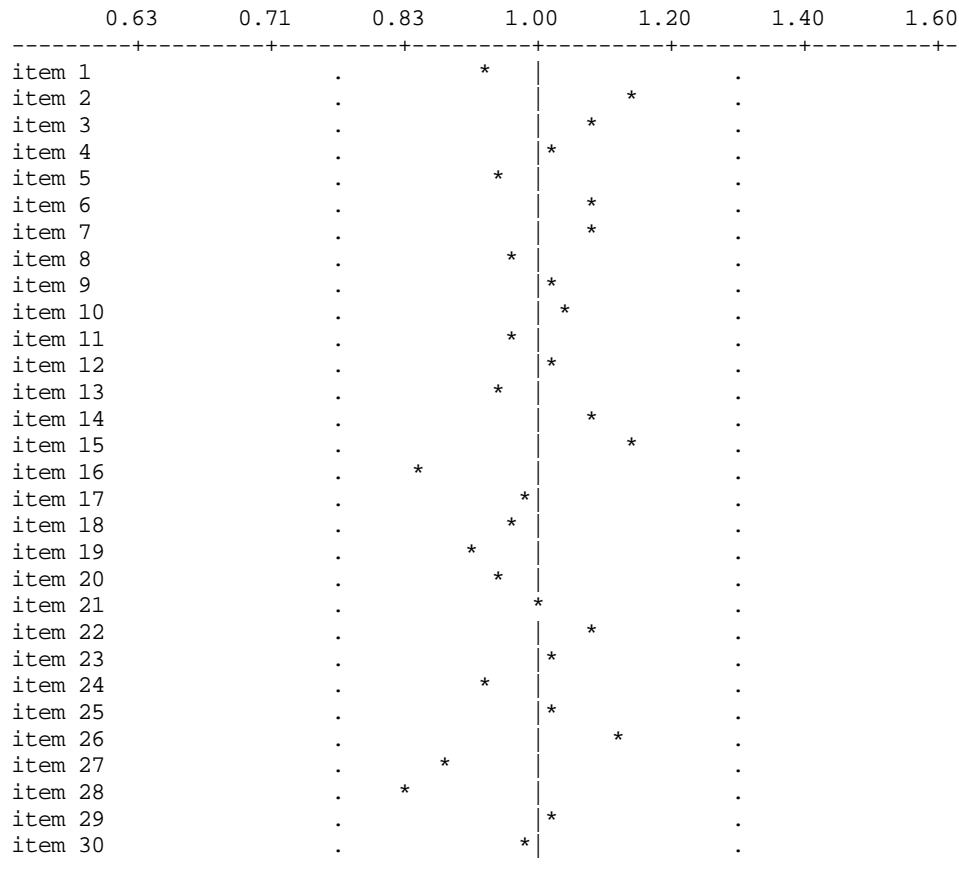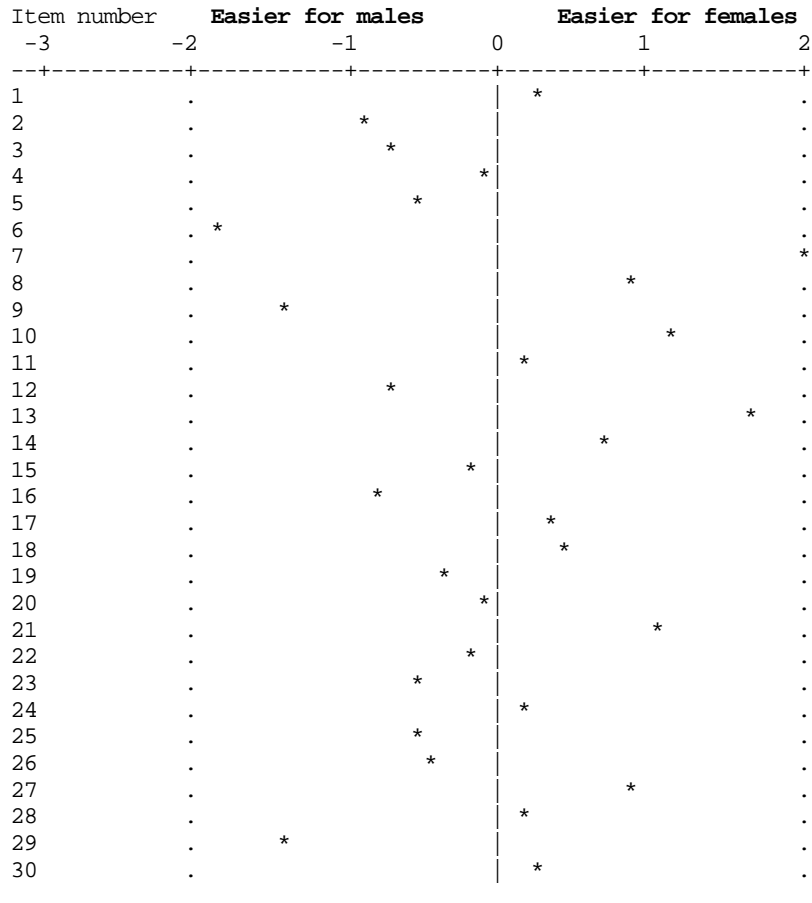
Figure 4: Plot of Infit Mean Squares for all Items

Figure 5 identifies only one question that is biased, Item 7, which contrary to the conventional wisdom is biased in favour of female students. An examination of the question revealed no obvious cause of concern, such as sexist language or a non-inclusive context. Rather, it is a straightforward question that requires the careful calculation of an answer using a formula presented in the text and discussed in lectures. Herein lies a possible source of bias. Female students have a tendency to be neater and more careful in the setting out of arithmetic questions, which in turn leads to greater accuracy and hence higher performance. The concept under review in the question is central to an important topic covered in the subject. On this basis, the question cannot be dropped from the examination in future years. However, if the source of the bias related to language or content then the question should be either reworked or dropped in the future.

Guessing

In the author's experience as both a student and lecturer in economics, multiple-choice questions are often synonymous with guessing. Hence, the emergence of the term multiple-guess exams and the development of punitive marking regimes where students tend to be punished for selecting the wrong answer. For example, a mark for a correct answer and minus one mark for an incorrect answer is not an uncommon practice. The author has even been involved in teaching a subject where the results of a multiple-choice test were scored using a procedure which penalised students for the possibility that they may have guessed the correct answer (Keats and Lord 1962). In such a learning environment the documented differences in the ways that identifiable groups of students

approach multiple-choice questions is ignored. For example, female students tend not to guess. Hence, they tend to be disadvantaged when there is no penalty for wrong answers and benefit when wrong answers are penalised (Kukhele, Thissen, and Wainer, 1994). These approaches are obviously based on the assumption that all students guess and that students guess between all possible distractors, rather than use some cognitive process based on the learning they have conducted to eliminate 'wrong' answers in order to reduce the number of distractors to choose between.

```
Item number    Easier for males           Easier for females
   -3          -2          -1          0           1           2
--+----------+-----------+----------+----------+-----------+
  1           .                      |    *                 .
  2           .               *      |                      .
  3           .                 *    |                      .
  4           .                    * |                      .
  5           .                 *    |                      .
  6           . *                    |                      .
  7           .                      |                      *
  8           .                      |       *              .
  9           .        *             |                      .
 10           .                      |            *         .
 11           .                      | *                    .
 12           .               *      |                      .
 13           .                      |                 *    .
 14           .                      |     *                .
 15           .                    * |                      .
 16           .              *       |                      .
 17           .                      |   *                  .
 18           .                      |    *                 .
 19           .                   *  |                      .
 20           .                    * |                      .
 21           .                      |         *            .
 22           .                    * |                      .
 23           .                *     |                      .
 24           .                      | *                    .
 25           .                *     |                      .
 26           .                 *    |                      .
 27           .                      |       *              .
 28           .                      | *                    .
 29           .        *             |                      .
 30           .                      |  *                   .
```

Figure 5: Plot of Standardised Differences

Quest provides information that can be used to determine the level of guessing in a multiple-choice test. The Outfit t statistic measures the degree of fit between the performance of individual students and the Rasch Model. The Outfit t compares the ability of students with the difficulty of the questions that they answered. Negative values indicate that students of higher ability are getting easy questions wrong, while positive values indicate that weak students are getting difficult questions correct. A value greater than positive 2 is the threshold for identifying students who are consistently guessing correct answers.

Analysis of the Outfit t statistics for all 844 students reveal that only 32 or 3.8 per cent of the class were consistently guessing, which is less than the five per cent allowed in the t-test. Of this small number of cases, only one student passed the test and two other improved their grades from a pass to a credit on the basis of the questions they guessed. However, the other 29 students had clearly failed the test, hence their performance was not significantly improved by guessing. On the basis of this evidence, it seems to be rather unfair to penalise all, if any, students for guessing.

Quest also facilitates the analysis of the performance of every student in order to determine the pattern of guessing on a multiple-choice test. Figure 6 is an example of a 'Kidmap' that can be obtained from Quest. Candidate 167 is an example of a weak student who scored 8 out of a possible 30. The Kidmap shows that five of this student's correct answers were greater than the ability of the student. Indeed, the student correctly answered the most difficult question.

This discussion provides evidence that guessing does occur on multiple-choice test, but not to the extent that multiple-choice should be a euphemism for multiple-guess. This study demonstrates that guessing is largely confined to weaker students whose performance is not improved by guessing.

```
Harder Achieved                                   Harder Not Achieved
                                  |   |
                15                |   |
                                  |   |7(4)
                                  |   |2(2)
                 9                |   |14(4) 23(3)
                                  |   |11(1) 12(1) 22(2)
                                  |   |6(1) 28(2)
                                  |   |26(1)
                13                |   |10(5) 17(1) 19(3) 30(1)
                                  |   |25(2)
                 5                |   |
                 8                |   |24(5) 27(3) 29(4)
                                  |   |20(5) 21(2)
                                  |   |4(1)
                                  |   |3(1)
                16                |   |(student ability)
                18                |   |
                 1                |   |
Easier Achieved                                   Easier Not Achieved
```

Figure 6: Kidmap Showing the Performance of an Individual Student

CONCLUSIONS

The resource constraints experienced by the Division means that multiple-choice questions are set to play an increasingly important role in assessment procedures in the large core subjects. Subject coordinators cannot afford the luxury of setting their own multiple-choice tests and are increasingly relying on the much cheaper and quicker method of tailoring suitable questions that are available in test banks and instructor's manuals. The source of these questions raises considerable concerns about their validity and reliability. To some extent this study allays these concerns as only one question failed to attain the minimum discrimination threshold and one was biased in favour of female students. However, a number of questions need to be modified to improve their discrimination power.

The study does not support the widely held view that students guess from all possible distractors if they do not know the correct answer. This study supports the view that students approach multiple-choice tests in much the same way that they approach short answer questions. That is, they attempt to answer the question and then look for the distractor that most closely fits they answer they obtained. The results imply that students should not be penalised for incorrect answers. Rather they should be awarded partial credit in recognition of the cognitive processes that lead to the choice of certain incorrect distractors.

REFERENCES

Adams, R.J. and Khoo S-T. (1993) *Conquest: The Interactive Test Analysis System,* ACER Press, Canberra.

Andrich, D. (1988) *Rasch Models for Measurement*, Sage, Beverly Hills.

Cole, N.S. and Moss, P.A. (1989) Bias in Test Use, in R. Linn, (ed.) *Educational Measurement, 3rd ed.*, MacMillan, New York, 201-219.

Hambleton, R.K. (1989) Principles of Selected Applications of Item Response Theory, in R. Linn, (ed.) *Educational Measurement, 3rd ed.*, MacMillan, New York, 147-200.

Keats, J.A. and Lord, F.M. (1962) A Theoretical Distribution of Mental Scores, *Psychometrika*, 27, 59-72.

Kukhele, R., Thissen, D. and Wainer, H. (1994) On the Relative Value of Multiple-Choice, Constructed Response, and Examinee-Selected Items on Two Achievement Tests, *Journal of Educational Measurement*, 31 (3), 234-250.

Rasch, G. (1960) *Probabilistic Models for Some Intelligence and Attainment Tests*, University of Chicago Press, Chicago.

Sheridan, B., Andrich, D. and Luo, G. (1997) *RUMM User's Guide*, RUMM Laboratory, Perth.

Snyder, S. and Sheehan, R. (1992) The Rasch Measurement Model: An Introduction, *Journal of Early Intervention*, 16 (1), 87-95.

Walstead, W.B. and Becker, W.E. (1994) Achievement Differences on Multiple-Choice and Essay Tests in Economics, *American Economic Review*, 84 (2), 193-196.

Walstead, W.B. and Robson, D. (1997) Differential Item Functioning and Male-Female Differences on Multiple-Choice Tests in Economics, *The Journal of Economic Education*, 28 (2), 155-169.

Weiss, D.J. and Yoes, M.E. (1991) Item Response Theory, in R.K. Hambleton and J.N. Zaal (eds) *Advances in Educational and Psychological Testing and Applications*, Kluwer, Boston, 69-96.

Master concept or defensive rhetoric: Evaluating Australian VET policy against past practice and current international principles of Lifelong Learning

Robin Ryan
Flinders University Institute of International Education *robin.ryan@flinders.edu.au*

Lifelong learning has become a frequently repeated mantra of national planning and policy agencies in Australian Vocational Education and Training (VET). Critics claim the use of the term is rhetorical rather than a real commitment to the principles and practices of lifelong learning. A review of research indicates that real progress has been made in implementing some aspects of lifelong learning in VET, but that contemporary practice falls far short of internationally agreed principles of lifelong learning or even the sector's earlier experience with the similar concept of recurrent education. Current policy settings impose too many barriers to the adoption of lifelong learning in VET and need comprehensive reform.

Vocational education, Education policy, Work-based education, Lifelong learning, Knowledge economy

INTRODUCTION: POLICY AND RHETORIC

Lifelong learning has gained wide acceptance in many countries as a 'master concept' for planning and debating all areas of education. It has a particular resonance for Australian vocational education and training (VET) since VET, along with the adult community education (ACE) sector,

- provides the educational opportunities which most broadly apply across age and socio-economic groups;

- delivers its programs through the widest variety of locations and means - regional, remote and urban, on-line and in class, in workplaces, in colleges, in community centres, in prisons; and

- accommodates the widest range of previous educational backgrounds.

Not surprisingly, lifelong learning has become part of the standard rhetoric of the sector, not least of its primary policy arm, the Australian National Training Authority (ANTA), although that body came relatively late to the idea (ANTA [1998b] seems to be the first, passing, reference). While ANTA's verbal endorsement has grown stronger as its own market research has shown how widespread enthusiasm for learning is throughout Australian society (ANTA 1999), it is not clear that this has led to any realignment of policy direction.

To describe the terminology employed in setting educational policy as rhetoric is not in itself a critique, nor is it intended to ascribe bad faith. The use of rhetoric is integral to the framing of educational policy debate. As pioneer policy theorists March and Olsen pointed out "the history of administrative reorganisation in the twentieth century is a history of rhetoric" (March and Olsen, 1983, p282; see also Radin and Hawley, 1988).

The use of rhetoric in policy debate is seldom to persuade, but more to set terms of engagement and criteria of success, to determine the roster of participants and to seek allies from distant but more influential policy communities (Buamgartner, 1989).

This pattern was established in the debate leading to the creation of ANTA, as States and Commonwealth sought to frame a debate about education and training in terms of a wider discussion of federal-State finance and the future of the federal governance structure (Ryan, 1999a). It is suggested that the pattern is now replicated by ANTA itself as it seeks to defend itself against an emerging critique. That critique argues that ANTA policy over-preferences business interests, devalues the role of individual students and communities, replaces educational goals with a narrow doctrine of vocational competency and trusts naively in a contrived market mechanism to produce efficient and equitable outcomes (see, for example, Billett and Hayes, 1998; Gonzci, 1998; Ryan, 1995; Fooks et al, 1997).

In short, it is a critique that suggests that training, in its narrowest and most utilitarian guise, has overtaken education or equity as the driving force of Australian VET. A potentially effective counter to this critique would be to claim that, on the contrary, Australian VET policy is firmly established within the international consensus represented by the lifelong learning philosophy. ANTA has attempted to make this claim: for example, by utilising the term in its national plan for VET (ANTA, 1999); by redefining an exercise in market research as the foundation for a lifelong learning strategy (Scollay, 2000); and by attempting to reinterpret its training regulatory framework in the now fashionable debate on a national innovation strategy (Scollay, 2001).

One unusual aspect of the current mobilisation of rhetoric is that it ignores entirely the VET sector's previous substantial experience with the cognate concept of recurrent or lifelong education. This paper seeks to evaluate the claim that national policy in Australian vocational education and training reflects the adoption of lifelong learning as its guiding principle. It seeks to compare the policy settings now in place in VET both with the earlier experience of lifelong education in TAFE and also with what can be established as an international consensus on the principles essential to effective strategies of lifelong learning.

It is argued that real progress has been made on a number of dimensions relevant to lifelong learning goals, but that the current policy agenda in many ways rates poorly against previous practice and fails most of the essential criteria of the international consensus. It is suggested that, by adopting a wholly utilitarian and economically focused agenda, ANTA ignores its own and other research on the motivation to learn (ANTA 1999; Golding and Volkoff, 1998) and, judged by international experience, is unlikely to attain even these narrow goals.

LIFELONG LEARNING: THEN AND NOW

Ideas about the need to learn throughout life are far from new. Comenius talked about learning from 'cradle to grave' in the seventeenth century (Ralph, 1999) while the French revolutionaries introduced a law for lifelong education into the National Assembly in their Year Two (Matheson and Matheson, 1996). There was a widespread movement for the continuing education of adults throughout Scandinavia and Western Europe in the nineteenth century (Kallen, 1996) and much the same spirit drove the Working Men's College movement in Australia (Murray Smith, 1996).

However, it was in the mid-1960s that the immediate antecedents of the current model of lifelong learning emerged in discussion in Europe on ideas variously referred to as lifelong education, recurrent education and *education permanente*. A compendium of these discussions was published at the end of the decade (Council of Europe, 1970) but the first statement to gain world wide prominence was the work of Faure and his colleagues on the UNESCO International Commission on the development of Education. This was published as *Learning to Be* in 1972 (Faure, 1972).

LIFELONG AND RECURRENT EDUCATION

Learning to Be was a radical document, echoing the sentiments on 'education as liberation' developed by Friere, but its basic theme had considerable appeal to Australian technical educators, as did its ringing demand that

> Rigid distinctions between types of teaching – general, scientific, technical and professional – must be dropped and education, from primary and secondary levels, must become theoretical, technological, practical and manual at the same time (Faure, 1972, p195).

At the same time, another vision of lifelong education was promoted by the Organisation for Economic Cooperation and Development under the title recurrent education (OECD, 1973). While supporting the UNESCO argument, the OECD had a firmer eye on the workplace and was especially concerned at overcoming the boundaries between the worlds of learning and of work. It was concerned at the persistence of distinctions between general and vocational education and, in the case of the latter, the use of traditional educational formats which failed to produce qualities of self-awareness and autonomous decision making in the learner (OECD, 1973, 43).

Australian technical educators were early adopters of the new ideas from Europe. The first to argue publicly for the use of lifelong education as a 'master concept' for planning vocational and adult education seems to have been Max Bone, the inaugural Director of Further Education in South Australia, in the Patricia Chomley Oration in 1972 (Bone, 1972). This echoed ideas, though not terminology, used by the Karmel Report on South Australian education (Karmel, 1969). At around the same time references to the new ideas began to appear in both Schools Commission and Advanced Education Commission reports (eg Advisory Committee on Advanced Education, 1969).

However, the greatest single influence in developing a practical application of the concept of lifelong education came from its enthusiastic acceptance by Kangan, the Commonwealth official whose 1974 report *Technical and Further Education in Australia* effectively created a new sector of education in Australia and empowered it with a strong philosophical base (ACOTAFE, 1974; Goozee, 1995).

The idea of lifelong education, usually in its OECD guise of recurrent education, became so identified with the TAFE sector that by 1977 a Chair of the TAFE Commission could speak of a diversion of resources to other education sectors as a dilution of the concept (Coughlan, 1978). Interestingly, the concept was welcomed with almost equal warmth by industry bodies interested in training, who saw it as breaking down barriers such as age restrictions on apprentice training (CAI, 1978).

Although recurrent education arose in the expansive era of the Whitlam Government, it was if anything more firmly embraced by its Coalition successor. The Fraser Administration used it to justify a string of equity and, indeed, social engineering policies in the TAFE sector, to the point that TAFE Directors warned of system overload (Ryan, 1999b; Goozee, 1995).

All this came to a sudden end in 1987 when TAFE was removed from the overview of the Commonwealth Tertiary Education Commission and placed in a new economically oriented Ministry of Education, Employment and Training, with Hon J. S. Dawkins as Minister. Even before assuming the Ministry, Dawkins had written the preface to a major report on industry training in terms whose emphasis on the needs of the labour market rather than the needs of the individual and the coining of the term 'skills formation' to replace TAFE and training, gave a

strong indication of the new directions which were to be taken in applying economic rationalism to vocational education and training (Goozee, 1995, p107).

On becoming Minister, Dawkins used resource agreements with the States to redirect funding from general and equity based education to vocational training and utilised a series of reports by federal agencies to replace educational goals in vocational education with a much more restricted notion of competency levels linked to industrial awards (Ryan and Hardcastle, 1996).

A further Commonwealth goal, originally set out in an Employment Department submission in 1985, was also acted on:

> that vocational education was not a community service but a training market, which like all markets, would work more efficiently with greater competition (Ryan and Hardcastle, 1996, p241).

In short, within a few years the philosophy of lifelong education which had underpinned Australian VET since 1974 was totally eliminated and replaced by a managerialist agenda which proclaimed its predecessor doctrine 'right for its times' but now as outdated as any other fashion from the 1970s. As the then Prime Minister argued:

> …we are moving beyond the Kangan report. The report was, of course, a product of its times. It was written in the context of a relatively stable manufacturing sector, sheltered behind a series of tariff walls and other sectoral support measures…All that has changed (Keating, 1994).

GLOBALISATION AND THE CHANGING ENVIRONMENT OF WORK

The new philosophy of vocational education and training, which has enjoyed bipartisan political support since 1987, undoubtedly has an internally clear and coherent rationale, which its advocates believe suited to the needs of the contemporary economy. However, as it is based on an explicit rejection of the principles of lifelong education in its 1960s and 1970s formulation, it is far from clear that it can be adapted to fit the 1990s growing interest in lifelong learning.

On the other hand, it must be accepted that lifelong learning has itself evolved considerably from its predecessors, while the world of work has equally changed in dramatic and largely unanticipated directions.

The new model of lifelong learning

A growing awareness of the consequences of globalisation lies behind the change from lifelong or recurrent education to lifelong learning. Lifelong learning differs from the earlier ideas because of changes wrought by the forces of globalisation, new technology and modern forms of the organisation of work. (Delors, 1996, 51).

The OECD education ministers, at a 1996 meeting chaired by the then Australian Minister (Hon Simon Crean), placed globalisation at the forefront of the factors impacting on education since their previous meeting in 1990. Noting that globalisation is by no means a new phenomenon, the ministers never the less remarked how,

> the concept has been broadened to encompass not only movements of goods and services, but also of investment, people and ideas across national and regional frontiers. Since the 1970s, three closely related phenomena have played a central role in facilitating and spurring a new wave of globalisation: market deregulation, the advent

and spread of new information technologies based on microelectronics, and the globalisation of financial markets (OECD, 1996a, 29).

UNESCO, in the Delors report, based its concept of lifelong learning on four pillars:
- Learning to Be
- Learning to Know
- Learning to Do
- Learning to Live Together (Delors, 1996)

The OECD saw lifelong learning as based on three fundamental objectives:
- Personal development,
- Social cohesion,
- Economic growth. (OECD, 1996a)

It argued that none of these can be taken in isolation. Lifelong learning must contribute to an array of aims rather than to a single goal.

In setting out their view, the OECD contrasted lifelong learning with its earlier formulation, recurrent education.

- The earlier model implied episodes of education between episodes of work or other activity. Now education is seen as continuous and embedded in the work experience.

- The last 20 years have seen a retreat of government from full support of education and alternative financing models have to be developed.

- Very high school retention rates and tertiary participation rates have now been achieved, changing the focus to intersectoral flows and equitable participation.

- Social demand is being replaced by individual demand (OECD, 1996a).

It is reasonable, then, to expect the policy settings for lifelong learning to differ from those applied to lifelong or recurrent education. Moreover, the world in which policy operates has changes dramatically over the last two decades.

CHANGING WORK, CHANGING SOCIETY

Most commentators agree with the OECD ministers that globalisation is the primary driver of international initiatives in lifelong learning. Globalisation is associated with rapid technological change, especially in information technologies, the transition from an industrial and service economy to a knowledge based economy, new ways of organising work and new attitudes to leisure and work (GLS, 1999; Cresson, 1999).

The new economy and the new workforce

In unpacking the components of globalisation, which call for lifelong learning as a response, commentators divide between optimistic and pessimistic scenarios.

Optimists note that knowledge is becoming the key factor in economic success and that in the new economy a workforce capable of rapid learning and innovation is critical. In the new economy, successful enterprises are those which commit to becoming knowledge organisations, in which customers, employees and managers constitute learning networks. Enterprises aim to become high performance workplaces and personal mastery is the goal of individual workers (GLS 1999; OECD 1996b). The skills workers need are not merely technical, but increasingly generic skills of communication, working together, analysis and problem solving (Robinson, 1999).

The pessimistic view argues that, while there is some evidence of the emergence of knowledge organisations (Harris & Volet, 1997), what has been more commonly found in Australia is an undervaluing of intellectual capital (Hopkins, 1998). Maglen and Shah concluded that the impact of globalisation, rapid technological change and economic restructuring on Australia has been more negative than positive (Robinson, 1999).

The pessimistic scenario is backed by evidence on changes to the nature and structure of the Australian workforce. While persistently high levels of unemployment have been reduced in recent years, only a bare majority of Australian workers enjoy traditional employment conditions of full-time, continuing tenure work with a single employer. Instead, the following apply:

- Underemployment and marginal workforce attachment are prevalent (Waterhouse et al, 1999) while those in employment experience work intensification (ACIRRT, 1999).

- Non-standard forms of workforce organisation- part-time, casual, outsourced, the use of labour hire companies- are increasingly common (ACIRRT, 1999; Burke, 1998; Robinson, 1999).

- New ways of organising work, separation into core and periphery employees (Waterhouse et al, 1997) and the emergence of the disposable workforce (Drago, 1999), are increasing.

Many of these characteristics of the new workforce are associated with lesser availability of employer supported training (Van Den Huebel & Wooden, 1999), while workplace training opportunities go most frequently to those with prior educational benefits (Fryer, 1997; Cresson, 1996; Watson, 1999b).

None of these developments reduces the need for lifelong learning. A nation's primary assets in a global economy are the skills and insights of its citizens (Reich, 1991). In some cases these are fostered by learning organisations committed to the mutual development of the organisation and its people (Senge, 1990). Alternatively, the growth of outsourcing, internal tendering, casualisation and short-term contracts shift responsibility for skill formation and maintenance from employers to individuals (Robinson, 1999).

Passing much of the risk, cost and responsibility of learning to individual workers means that the potential for lifelong learning to develop in an equitable, socially inclusive and personally satisfying manner is significantly reduced.

Demography and lifelong learning

Most advanced economies are experiencing an aging of their populations and workforces.

In Australia, the number of 15-24 year olds is expected to remain constant in the next two decades despite population growth of 20 per cent, while the number in the 45-64 age group will grow by over 40 per cent (Robinson, 1999). This significant population shift means that policy must place as much attention on the learning needs of adults as of young people. With a constraint on immigration intakes, much of the source of new skills must come from the existing workforce. At the same time, a larger cohort of older citizens will make their own demands on the education and training system.

Australian education systems have not been well geared to the vocational and other learning needs of older people. A re-orientation of education and training, to shorter modules and different forms of student services, is required (Smith, A 1999). Funding and policy discrimination against the adult community education sector also need reform if lifelong learning is to have meaning for this large and growing segment of society (Watson, 1999b).

WHAT WOULD A LIFELONG LEARNING FUTURE LOOK LIKE?

Lifelong learning is the lifeblood of the knowledge economy. The evidence suggests that the new, knowledge economy is already a reality, although it overlaps with the residual, old economy.

- Trends in the major OECD economies show that knowledge based production already accounts for more than half of national product, that high technology industries are growing faster than manufacturing generally and that knowledge intensive service sectors are growing faster still (OECD, 1996b).

- The knowledge economy's emphasis on the diffusion and use of knowledge creates a 'network society' in which the social and economic success of individuals and firms depends on developing relationships, which promote learning (OECD, 1996b).

- The learning society is one which builds on these networks to develop an interactive, information-rich society in which learning is widely diffused, with a culture that values learning throughout life for all citizens (Kearns, 1999).

- Most European countries have favoured active public intervention rather than market forces alone to mould learning networks into desired channels, for example the SOCRATES and LEONARDO programs of the European Commission and the local learning consortia or Italy and the Netherlands (Papadopoulos, 1999).

Education and training in the knowledge economy needs both technical and 'soft' skills, but it is important that it be geared less to specific vocational competencies and more to the development of key or generic competencies transferable to a wide range of tasks (Moy, 1999).

Principles of lifelong learning

The concept of lifelong learning cannot be reduced to a single set of precepts, although most commentators would accept the core principles of UNESCO and OECD (Delors, 1996; OECD, 1996a). These have been built on to establish a substantial international consensus, represented by major documents published by public authorities in Britain, Japan, Finland, the Netherlands, the Council of Europe and the European Union (Papadopoulos, 1999; Cresson, 1996).

Thus, while there is no single model of lifelong learning equally applicable to all national and cultural contexts, there is a consensus on key strategies. These include:

- putting the individual learner and stressing quality and flexibility in learning within a coherent strategy embracing all forms of education (OECD, 1996a; Fryer, 1997);

- promoting the widest possible participation in education and training for all age groups, with emphasis on foundation learning and effective transitions (OECD, 1996a);

- the development of social partnerships and learning networks (Papadopoulos, 1999), including those which bring education and business closer (Cresson, 1996) while not restricting education to meeting the needs of business (OECD 1996a);.and

- priority to those most in need with special attention to adult learning, especially for literacy difficulties (OECD, 1996a).

Lifelong learning requires policies which:

- do not discriminate among sectors or on perceived vocational/non-vocational distinctions (SEETRC, 1997);

- build appropriate pathways not only from school to work and through the education sectors but for all needs of adult learners (Robinson, 1999); and

- minimise artificial barriers erected by accreditation rules, funding regimes and intergovernmental relations (West, 1998; NBEET, 1995).

Partnerships and networks are part of the innovation cycle in the knowledge economy. Lifelong learning will often be expressed through the creation of learning organisations, the development of professional, geographic and virtual learning communities, and even through the creation of learning cities (these ideas are explored extensively in Kearns, 1999).

For Australia there is a pressing need to rediscover the social, citizenship and equity dimensions of lifelong learning which are crucial to most overseas policy development (OECD, 1996a and 1998; Papdopoulos, 1999). Australian economic policy makers have conceded that competition policy and associated reforms incur costs in terms of social and regional impacts (Productivity Commission, 1999). A similar recognition of the risks and disadvantages of a wholly instrumental approach to education and training is overdue.

HOW FAR HAVE WE COME?

Australia has made significant progress in expanding participation in various forms of education and training.

- In the last decade, there has been enormous growth in higher education; on present participation patterns, 45 per cent of the population will undertake higher education at some stage in their lives

- Participation has grown most rapidly in the 20-24 age cohort; female enrolments now exceed male, although there is great gender segmentation by educational field.

- 16 per cent of the 20-24 cohort participates in higher education, while 17 per cent participates in VET; the latter has grown more slowly, but rates for females, indigenous people and those of non English speaking background have converged on male participation (Ball, 1999; Marginson, 1999)

- 20 per cent of the population aged over 15 years is enrolled in a formal course of training each year and 60 per cent of employers provide some kind of training for their employees each year (Robinson, 1999).

Looking only at the apprenticeship system, important changes have been achieved in policy which recurrent education in the 1970s challenged without significant success.

- Age restrictions have been decisively overcome.

- There has been some growth in part-time and school based apprenticeships.

- The occupational base covered by 'new' apprenticeships has significantly widened.

- There has been a major increase in female participation: although it still lags the proportion of females in the workforce, this is partly countered by the very much greater female than male participation in higher education.

- There has been a growth in indigenous and non-English speaking background apprentices.

- There has been an improving position for people with disabilities (NCVER, 2001).

Watson concluded from such indicators that "Australia appears well on the way to becoming a learning society" (1999a, p1).

However, there are indicators, which suggest Australia remains far from a learning society.

- Australia has recently experienced an actual decline in total participation in education and training (Watson, 1999a).

- Year 12 retention rates have been declining since 1992; Australia is the only OECD country to experience this (Ainley & McKenzie, 1998).

- Only 53 per cent of Australians in the 25-64 age band have completed Year 12, ranking Australia at 17 out of 25 OECD countries, and 19th for 25-34 year olds (McKenzie, 1998).

- There is now no prospect of achieving the Finn participation targets for 19 year olds and more than a quarter of 20-24 year olds (32 per cent of females) are neither in full-time employment nor studying for a recognised qualification (Ball, 1999).

- Australia is at best a mid-level performer in adult literacy: 45 per cent of adults will experience literacy difficulties at work and in community life (McKenzie, 1999).

Only limited progress has been made in the construction of intersectoral bridges and pathways.

- Although two thirds of VET students in the 20-24 age group have completed Year 12), only one in ten commencing university students does so from a background in VET.

- About 5 per cent of VET enrolments are of students with a completed higher education qualification, with a similar proportion having partially completed university courses.

- After decades of effort, transfers between universities and VET remain marginal to each sector (Ball, 1999; Marginson, 1999).

- There remain systemic difficulties in securing broad based credit transfer from VET to universities and the use of competency based criteria undermines university-VET articulation (NBEET, 1995).

- Adult and community education continues to be a vital part of learning in Australia, catering to around one million people each year, but artificial distinctions between vocational and non-vocational distort public policy on ACE (SEETRC, 1997).

- Although there has been recent growth in vocational education programs in schools, too much emphasis has been placed on relabelling to inflate participation numbers (Malley, 1999); VET in schools remains marginalised and is not well served by accreditation arrangements (Ryan, 1997).

The hope that a training, or even more, a learning culture is emerging in Australian enterprises has some support.

- Over 80 per cent of employees receive some kind of training from their employers, although the vast majority is unstructured, on-the-job training (Robinson, 1999).

- A study of 17 Australian companies identified common characteristics in which purposeful learning contributes to individual, team and organisational development (Harris & Volet, 1997).

Against this are substantial reasons for believing that a learning culture remains far from reality.

- Only 18 per cent of Australian companies provide structured training for their employees (Peoples, 1998).

- Employer expenditure on structured training decreased on all measures between 1993 and 1996 (Burke, 1998).

- Learning organisations remain far from common and the absence of a training, let alone a learning culture, in small enterprises is well documented (Kearns, 1999; Field, 1997).

- Australian businesses characteristically undervalue intellectual capital (Hopkins, 1998).

POLICY BARRIERS TO LIFELONG LEARNING IN THE VET SECTOR

Many VET policy stances have created outright barriers to the implementation of lifelong learning. Often this derives from an unbalanced, economic view of VET with an emphasis on short-term industry relevance, rather than the simultaneous pursuit of the array of personal, social and economic objectives advocated by the OECD (OECD, 1996a). For example, the emphasis on skill widening inherent in the New Apprenticeship system is to some degree at the expense of skill deepening. Traditional (Level III) trade and related apprentice numbers have largely recovered from their collapse of the early 1990s, but remain substantially below their peak in 1990 and are failing to grow at anything like the rate of lower level (I and II) training (NCVER, 2001).

Some commentators argue that there has been a decline in the quality of apprentice training, whether through the imposition of a mandated competencies regime or because of a lower ability recruitment pool (Smith, L, 1999). Diploma courses in VET have failed to grow at anything like their counterparts in higher education (Ball, 1999).

The division between general or academic and vocational education remains as wide as ever and pathways between higher education and VET remain poorly constructed because of a separate accreditation system in the VET sector, derived from industrial awards (Sweet, 1993), which acts as a barrier to mutual recognition (NBEET, 1995).

Adult community education remains neglected and, in so far as it is covered by VET policy and funding, the vocationalist orientation acts to distort the goals of ACE (SCEETRC, 1997).

Development of social partnerships, apart from engagement with employer bodies, has largely been ignored and key stakeholders, including regional communities, complain that they are excluded from the policy process (Billett and Hayes, 1998).

While the need to focus on individual learners is at the forefront of the international debate on lifelong learning, individual students are rarely consulted by VET policy makers (Golding & Volkoff, 1998). VET policy assumes an identity of interest between students and their current employers, despite the contrary picture revealed in TAFE and VET graduate destinations surveys (eg ANTA, 1998).

CONCLUSION: GETTING SERIOUS ABOUT
LIFELONG LEARNING IN AUSTRALIAN VET

Despite the increasingly prevalent rhetoric, lifelong learning is not a firmly established goal of contemporary Australian VET policy. The theoretical base underlying a decade of VET reform is not easily compatible with principles of equitable lifelong learning (Gonzci, 1998).

This does not mean that the sector can simply revert to the policy framework of the TAFE 'golden age' of recurrent education in the 1970s. The challenge from technology and globalisation is sharper now. The rationale of western governments has changed, so that proposals for generous public funding and universal public provision in any education sector are unsustainable.

Even so, a strong case can be made that 1970s TAFE policy provided a firmer foundation for contemporary ideas of lifelong learning, as evident in international writing and practice, than does current VET policy. The task is to take what is valuable from the sector's previous experience and to apply it in terms compatible with present needs and realities.

A commitment to lifelong learning in VET would require policymakers to undertake a fundamental reappraisal of core components of present policy.

Four areas for reform are suggested.

1. Reshape VET policy to put the individual learner at the centre of education and training

The most striking characteristic of VET policy in Australia since the mid-1980s has been its indifference to the demands and needs of individual learners. Public policy has defined the client of VET as 'industry' or 'enterprises'. VET policy making has adopted a model in which learning is assumed to take place as a result of a training need agreed between an enterprise and an employee; students are often referred to as 'employees', as if this were the usual mode of VET learning.

This approach is contradicted by all available evidence. It is clear from graduate destination studies (ANTA, 1998) and from research on motivation for training (ANTA, 1999) that the official model applies only to a small minority of VET students. The vast majority of students are either school leavers seeking to gain a first job, or adults hoping to re-enter the employed workforce or to gain work with a different employer. Even in cases of planned enterprise training, the individual is most commonly the source of the decision to train (Smith & Hayton, 1999) and personal interest and self-fulfillment rank highly as motivators (Golding & Volkoff, 1998; ANTA 1999).

Although ANTA's recent marketing studies have reinforced the importance of a focus on individual learners, the response of the ANTA Ministerial Council has been once more to place marketing to industry ahead of marketing to the community (ANTA, 2000).

2. Recognise the diversity of stakeholders in VET and the need to build networks and partnerships

The greatest strength of a decade of VET reform has been an increased focus on relations between training providers and employers. The corresponding weakness has been a refusal to contemplate a wider range of stakeholders, especially the notion of learning communities. There is a strong feeling within Australia's regions that both individuals and communities have been the overlooked parties in VET (Billett & Hayes, 1998). The knowledge economy creates a network society where the ability to enter learning relationships is crucial for individuals and firms (OECD, 1996b). Most European governments have favoured active public intervention to mould learning networks (Papadopolous, 1999). Australian VET needs to recognise the social dimension of learning and the range of stakeholders who influence learners.

3. Refocus competencies on the career needs of individual learners, not on the past practices of industry.

Given the increasing transfer of responsibility for lifelong learning from governments and industry to individuals, it becomes less important to identify specific vocational competencies in existing industries. What is needed now is to ensure that all post-compulsory education becomes highly developed at imparting the generic skills that are increasingly important throughout the workforce, especially for new jobs emerging in the information age (Robinson, 1999). Frequently, these will be 'soft skills' identified in reports like *Enterprising Nation*, still largely ignored by VET and universities (Sheldrake, 1997).

4. Value the expertise of educators

Transference of specific competencies to generic skills does not happen easily. Industry endorsement of competencies is not a substitute for educational expertise in planning systematic learning experiences. Education and training programs must change continuously as skill requirements change in order to meet the diverse learning styles of adult learners (Robinson, 1999). One-size-fits-all solutions such as the Training Packages imposed by ANTA as a condition for funding and accreditation have limited utility in a lifelong learning philosophy.

The development of meta-learning ('learning to learn') skills for a population with a wide range of prior learning experiences is a major educational challenge. The disdain for educational expertise implicit in much public policy in the VET sector is a major barrier to success in lifelong learning.

5. Construct more flexible adult learning pathways

Pathways between VET and university education remain poorly developed (Marginson, 1999). While universities need to approach the issue of articulation more constructively, barriers constructed within VET should also be removed. Rigid systems of competency based curriculum impede intersectoral transfer (NBEET, 1995). More generally, pathways need to be developed which recognise the wide variety of adult learners, together with accrediting and recording systems which allow a diversity of learning experiences to be matched, when required, with formal qualifications (Robinson, 1999).

6. Rethink the role of competition and the way VET is funded

The Commonwealth-State review of the ANTA Agreement in 1996 warned against seeing competition as an end rather than as a means towards responsiveness and efficiency (Taylor, 1996). Research indicates that some elements of the competitive agenda have led to a weakening of quality in both public and private providers (Smith, L, 1999; Schofield 2000a, b, c).

The appropriate role of a competitive training market needs to be balanced against the benefits of collaboration and the risks to quality. A clearer definition of the role of the public provider, recommended in the Taylor review (1996), remains to be undertaken. Similarly, the best means of providing low demand programs, the needs of regional and remote communities and a broad range of equity issues remain unresolved.

FUNDING, PRIORITIES AND STRATEGIES

The contemporary lifelong learning literature recognises that, in contrast to 1970s views which saw public funding as the only option, private benefits from education and training need to be matched by household and enterprise contributions to its cost (OECD, 1996a). The funding issues raised in the West Review require addressing (West, 1998) so that funding arrangements facilitate

rather than hinder mobility between sectors (Robinson, 1999). Funding discrimination against adult community education should be avoided (SEETRC, 1997).

Increased public funding is almost certainly required by a policy of lifelong learning, but additional claims on the public purse need to be realistic and focused on achieving equitable outcomes.

There is a strong public policy case for providing those who have benefited least from formal education with a route back to lifelong learning (OECD, 1998). OECD estimates that costs of extending lifelong learning to adults who have not completed secondary education will entail an annual expenditure of three per cent of GDP, on top of education's current five per cent share (McKenzie, 1999). To include even more disadvantaged adults, those at the lowest literacy levels, would mean a doubling of education's share of GDP.

Thus the implementation of even the early stages of a lifelong learning philosophy will require rigorous prioritisation. Policy makers will need to explore innovative funding mechanisms, balancing contributions from individuals, industry and government. There is a need to discard traditional debates in which universal public funding or the simple application of market forces are seen as the only choices.

Beyond the vitally important resource issues, policy makers need the nerve and imagination to admit where their existing course is faulty and to envision a new way. Some of the new strategy might look like elements of the policy abandoned in 1987, but other elements will be entirely new. What is needed is not a return to the policies of the golden age, any more than the present international consensus is simply a duplication of earlier models. The need is for new strategies as imaginative as those developed by pioneers like Kangan, Karmel, Bone and their contemporaries. The international community's experimentation with lifelong learning provides a wide menu of choices.

REFERENCES

ACIRRT, (Australian Centre for Industrial Relations Research) 1999. *Australia at Work.* Sydney: Prentice Hall.

ACOTAFE (Australian Committee on Technical and Further Education) 1974. *TAFE in Australia*. Canberra: AGPS.

Advisory Committeee on Advanced Education, 1969. *Report for the 1970-72 Triennium*. Canberra: AGPS.

Ainley J and P McKenzie 1998. *The Influence of School Factors in Australia's Young Adults: the deepening divide*. Sydney: Dusseldorp Skills Foundation.

ANTA, (Australian National Training Authority) 1998a. *TAFE Graduate Destinations Survey, 1998*. Adelaide: NCVER.

ANTA, 1998b. *A Bridge to the Future*. Brisbane.

ANTA, 1999. *National Marketing Strategy for Skills and Lifelong Learning*. Brisbane.

Ball, K 1999. Participation in and Outcomes from Vocational Education and Training in *Australia's Young Adults: the deepening divide*. Sydney: Dusseldorp Skills Foundation.

Baumgartner, F 1989. *Conflict and Rhetoric in French Policy Making*. Pittsburg: University of Pittsburg Press.

Billett, S and S Hayes1998. *Meeting the Demand: the needs of TAFE clients*. Adelaide: NCVER.

Bone, M H 1972. *Lifelong Education: the Role of Permanent Education in the Education System.* Adelaide: Australian College of Nursing.

Burke, G 1998. Economic Change: *Implications for Jobs, Income and Joblessness.* Melbourne: CEET Conference on Rapid Economic Change and Lifelong Learning.

CAI (Confederation of Australian Industry) 1978. *Education and Training in Australia*. Melbourne.

Coughlan, K 1978. Opening Remarks in C. Duke (ed*) Recurrent Education 1977: Trends, Tensions, Trade Offs*. Canberra: AGPS.

Council of Europe 1970. *Lifelong Education: A Compendium of Discussions*. Strasbourg.

Cresson, E 1996. Towards a Policy of Lifelong Learning. *Vocational Training European Journal*, 8/9.

Delors, J (Chair) 1996. *Learning: the Treasure Within*. Paris: UNESCO.

Drago R 1999. Workplace Transformation and the Disposable Workplace: employee involvement in Australia. *Industrial relations*, 35 (4).

Faure, E 1972. *Learning to Be*. Paris: UNESCO.

Field, L 1997. *Training and Learning in Small Business*. Sydney: Research Centre for Vocational Education and Training, UTS.

Fooks, D, R J Ryan and K Schofield 1997. *Making TAFE Competitive*. Canberra: Australian College of Education.

Fryer R H 1997. *Learning for the 21st Century*. London: First Report of the National Advisory Group for Continuing Education and Lifelong Learning.

Golding B and V Volkoff, 1998. Participation in VET in Australia; different drives, same needs in Ferrier, F and D Anderson (eds) *Different Drums, One Beat?* Adelaide: NCVER.

Gonzci, A 1998. The Potential Destruction of the Vocational Education and Training System in Ferrier, F and D Anderson (eds) *Different Drums, One Beat*? Adelaide: NCVER.

Goozee, G 1995. *The Development of Technical and Further Education in Australia*. Adelaide: NCVER.

Harris L and S Volet, 1997. *Developing a Learning Culture in the Workplace*. Perth: Murdoch University.

Hopkins S 1998. *The Role of the Enterprise in Lifelong Learning: what can be realistically expected?* Melbourne: CEET Conference on Rapid Economic Change and Lifelong Learning.

Kallen D 1996. Lifelong Learning and Retrospect. *Vocational Training European Journal*, 8/9.

Karmel, P 1971. *Education in South Australia*. Adelaide: Government Printer.

Kearns P 1999. *VET and the Learning Age*. Report by Global Learning Systems. Adelaide: NCVER.

Keating, P 1994. Foreword in Hall, W and P Kearns, (eds) *Kangan: 20 years On*. Adelaide: NCVER.

Malley, J 1999 *Implementation of Vocational Education and Training in Australian Secondary Schools*. Canberra: Australian Vocational Education Research Association Annual Conference.

March J G and J P Olsen 1983. Organizing Political Life: What Administrative Reorganisation Tells Us About Government. *American Political Science Review*, 77.

Marginson, S 1999. Young adults in Higher Education *in Australia's Young Adults: the deepening divide*. Sydney: Dusseldorp Skills Formation.

Matheson D and C Matheson, 1996. Lifelong Learning and Lifelong Education: a critique. *Research in Post Compulsory Education*, 1 (2).

Moy J 1999. *The Impact of Generic Competencies on Workplace Performance*. Adelaide: NCVER.

Murray Smith, S 1966. Technical Education: the lines of development in C Sanders (ed) *Education for Development*. Perth: University of Western Australia Press.

NBEET (National Board of Employment, Education and Training) 1995. *Cross Sectoral Collaboration in Post Secondary Education and Training*. Canberra: AGPS

NCVER, (National Centre for Vocational Education Research) 2001. *Apprenticeships: Research at a Glance*. Adelaide.

OECD (Organisation for Economic Cooperation and Development) 1973. *Recurrent Education*. Paris.

OECD, 1996a. *Lifelong Learning for All*. Paris.

OECD, 1996b. The Knowledge Based Economy. *OECD Working Papers IV* (50). Paris.

OECD, 1998. *Education Policy Analysis*. Paris.

Papadopoulos, G,1999. Policies for Lifelong learning: An Overview of International Trends in *VET in the Learning Age*. Adelaide: NCVER.

Peoples, K 1998. Who Will Stand with the Young? Melbourne: Australian Education Union.

Productivity Commission, 1999. *Impact of Competition Policy Reforms on Rural and Regional Australia*. Canberra: draft report.

Radin B A and W D Hawley, 1988. *The Politics of Federal Reorganisation*. New York: Pergamon.

Ralph D 1999. *Lifelong Learning: Panacea or Pipedream*? Darwin. ACE/ACEA Joint National Conference.

Reich, R 1991. *The Work of Nations*. New York: A A Knopf.

Robinson, C 1999. *NewSkills, New Pathways: Lifelong Learning* is the Key. Canberra; University of Canberra Conference on Post Compulsory Education and Training – Looking to the Future.

Ryan R J and L Hardcastle, 1996. Technical and Further Education in Parkin, A (ed) *South Australia, Federalism and Public Policy*. Canberra: Federalism Research Centre, ANU.

Ryan, R J 1997. *VET in Schools*. Adelaide: NCVER.

Ryan, R J 1999a. How TAFE Became Unresponsive: A Study of Rhetoric as a Tool of Education Policy, *ANZ Journal of Vocational Education Research*, 7 (2.

Ryan, 1999b. From Recurrent Education to Lifelong Learning in *Report of the Invitational Seminar on the Delors Report*. Adelaide: Flinders University Institute of International Education.

Schofield, K 1999. *A Risky Business: Review of the Quality of Tasmania's Traineeship System*. Hobart: Office of Vocational Education and Training.

Schofield, K 2000. *Delivering Quality*. Melbourne: Office of Post Compulsory Education, Training and Employment.

Scollay, M 2000. Winds of Change Refresh the Learner's World. *Australian Training Review*, 36.

Scollay, M 2001. Innovation. *Australian Training Review*, 37.

SEETRC, (Senate Employment, Education and Training References Committee) 1997. *Beyond Cinderella*. Canberra: the Senate.

Senge, P 1990. *The Fifth Discipline: the art and practice of the learning organisatio*n. New York: Doubleday.

Sheldrake, P 1997. *Great Expectations: education and the world of work*. Adelaide: A W Jones Lecture, Australian College of Education.

Smith, A, 1999. *Creating a Future*. Adelaide: NCVER.

Smith, A and G Hayton (1999). What Drives Enterprise Training? Evidence from Australia. *International Journal of Human Resource Management*. 10(2).

Smith, L, 1999. *The Impact of User Choice on the Queensland Training Market: a progress report*. Brisbane: Department of Employment, Training and Industrial Relations.

Sweet, R 1993. *A Client Focused Vocational Education and Training System*? Sydney: Dusseldorp Skills Forum.

Taylor, R 1996. *Review of the ANTA Agreement*. Canberra; AGPS.

Van Den Huebel A and M Wooden, 1999. *Casualisation and Outsourcing: Trends and Implications for Work Related Training*. Adelaide: NCVER.

Waterhouse, P, B Wilson and P Ewer 1999. *The Changing Nature and Patterns of Work and Implications for VET*. Adelaide: NCVER.

Watson L 1999a. *Lifelong Learning for All: a policy challenge for government*. Toowoomba: Vocational Education and Training Research Conference.

Watson L 1999b. *Lifelong Learning in Australia; analysis and prospects*. http://www.Network.canberra.edu.au/lll/PUBLICATIONS/DiscussionPaperNo1final.htm

West R (Chair) 1998. *Learning for Life: report of the review of higher education funding and policy*. Canberra.

Looking backward, looking forward: The first 18 months of local governance and management in South Australia

Paul Kilvert
Director Partnerships 21 Taskforce, S.A. Department of Education, Training and Employment

Helen Tunbridge
Partnerships 21 Taskforce Superintendents

Alastair Dow
Partnerships 21 Taskforce Superintendents

The implementation of Partnerships 21 precipitated wide-ranging and profound changes in the policy, organisational and cultural dimensions of the public education system in South Australia.. In particular two state wide surveys were conducted at the end of the first 6 months of its operation and then at the end of the first year. The results of these two surveys were reported back to all sites and it is the opinions of site leaders expressed in the resulting documents that form the research basis of this paper.

Partnerships 21 Taskforce, South Australia, survey, school management, governance

LOOKING BACKWARD

In 1998 the South Australian Government commissioned a report on local management of public schools. The report "Community Partnerships in Education", known as the Cox Report after the chairperson of the review committee, Associate Professor Ian Cox, was presented in December 1998, following an extensive process of community consultation. Based on the recommendations of the report, the South Australian model of local management, named "Partnerships 21", was launched by the Premier of South Australia in April 1999.

In a feature unique to the South Australian model of local management, Partnerships 21 included both schools and preschools. Participation in the program was made voluntary and in August 1999 schools and preschools were invited to indicate their interest in joining the first round of Partnerships 21 for the start of 2000. 40% of all schools and preschools across the state chose to do so. By the start of 2001 a total of 75% have chosen to join the program.

The implementation of Partnerships 21 precipitated wide-ranging and profound changes in the policy, organisational and cultural dimensions of the public education system in South Australia. Partnerships 21 also coincided with two other major reforms - the review of the Education Act, and the review and rewriting of the department's curriculum framework. In relation to the former, the South Australian parliament in December 2000 passed amendments to the Education Act that provided the legislative basis for the function of governing councils under Partnerships 21. In relation to the latter, the department in 2000 developed the South Australian Curriculum, Standards and Accountability framework that is being implemented this year in all schools and preschools in the state system.

The impact of the changes have been deeply felt in the education department itself, as the balance of responsibility and accountability for the deployment of human resources, facilities and finance

shifted towards the schools and preschools[1]. This process of redefinition of responsibility and accountability has been evolutionary, and strongly shaped by the leaders of Partnerships 21 schools and preschools, leading to the review and redevelopment of systemic policies.

The development and implementation of the Partnerships 21 model was powerfully influenced by a series of 'policy shaping groups' that were established in 2000 to clarify and propose resolutions to policy issues affecting the operation of Partnerships 21 particularly as they applied to the global budget. Leaders from schools and preschools in country and metropolitan locations worked with leaders in state office to make up the membership of these groups in the areas of Curriculum Programs, Human Resources, Finance and Accounting, Physical Resources, Special Education, Structures and Governance, and Country Issues. Much of the work of these groups was directed to greater freedom and flexibility in local decision-making. The recommendations of the Policy Shaping Groups and the action taken by the Corporate Board in relation to each of them were summarised in the booklet, entitled "Departmental Response to the Recommendations of the Partnerships 21 Policy Shaping Groups", and circulated to all departmental sites. On the basis of these recommendations, the global budget structure was improved for 2001. The process and its outcome highlighted the nature of the changing relationship between local leaders and state office leaders through local management.

At another level, Partnerships 21 has been further shaped by the use of monitoring processes that involve the planned collection of data and information from P21 sites. In particular two state wide surveys were conducted at the end of the first 6 months of its operation and then at the end of the first year. The results of these two surveys were reported back to all sites in the two publications Partnerships 21 – from plans into practice, the first six months (August 2000), and Partnerships 21 in 2000, Looking Back, Looking Ahead (February 2001), both of which are available on the P21 website at www.partnerships21.sa.edu.au. The opinions of site leaders expressed in these documents form the research basis of this paper.

Partnerships 21 – the big ideas

Local operation of Partnerships 21 focuses on improving student learning outcomes through three significant areas of change, broadly termed *Partnerships*, *Quality Improvement* and *Resource Flexibility*.

Changes to *Partnerships* involved a redefinition of the authority and accountability of the school council or pre-school management committee, now collectively named 'governing council'. Partnerships 21, allied with the changes to the Education Act, has focused on strengthening the role of parents and the community through the governing council. In the schooling sector this has involved changes to the membership of the governing council, the adoption of a constitution, and most significantly, enhanced authority to approve and monitor the school's policies, strategic plan and budget. Each Partnerships 21 site's governing council is accountable to the Minister for Education and Children's Services for the strategic objectives and policies of the school/preschool outlined in its Partnerships Plan and for its delivery of statewide policies and programs outlined in the Services Agreement.

Quality improvement integrated the development of a three-year strategic plan, with performance indicators and improvement targets. While the concept of strategic planning is familiar to South Australian schools and preschools, the Partnerships Plan specifies measurable improvement targets and is linked to the site's three year resources strategy and Services Agreement. The

---

[1] Collectively, schools and preschools are termed 'sites' in this paper.

Services Agreement is one of the innovative elements of Partnerships 21, as it formalises the commitment of the department's state office to deliver a range of services to sites entering the program, as well as the sites' obligations to its community and state office. In this context, each site leader is accountable to the governing council for the implementation of the site's strategic plan and to the Chief Executive for the quality of its educational programs and outcomes.

Resource flexibility involved the establishment and operation of the global budget. A core principle of the Partnerships 21 model is that resources would be used more effectively when the local site makes the decisions about their use. This contrasts with the traditional model of centralised dispersal of funding through a series of grants and control of human resources through a centralised deployment and transfer system. The global budget allocates resources based on a new allocative mechanism using the following components: a base allocation for each site; a base allocation for each student; an additional allocation according to their year level; an additional allocation for students with special learning needs eg aboriginal students, students with disabilities, English as Second Language students, students from low socio-economic backgrounds.

The essential and consistent message at the centre of the Partnerships 21 model and its implementation is that the improvement of student learning outcomes will best occur through making effective connections between these three key elements.

Managing the change process

Following the release of the Cox Report, the department's Chief Executive, Mr. Geoff Spring, moved to establish a group of senior leaders from schools, preschools and district offices chosen for their capacity to contribute to what was to be a complex process of systemic change. Known as the 'Partnerships 21 Taskforce', the group was established in 1999 under the leadership of its director Dr Paul Kilvert. Its role has been to support local leaders in the change process and to facilitate the liaison between local sites and state office directorates as these directorates operationalise the recommendations of the Cox Report. The taskforce's work has focused on the coordination of training and development activities, the coordination of the policy shaping groups process, the publication of support materials in both written and electronic form, working with principals, directors and district superintendents in assisting individual sites, and significantly, working with officers in state office in conveying issues from local leaders and recommending solutions.

Central to the Partnerships 21 model has been its design as a comprehensive, integrated model of local governance and management. From the outset, the promotion of the model in the education community has emphasised the articulation of the key elements, with the primary objective of improving student learning. There is a clear awareness of the Partnerships 21 project as a multifaceted but conceptually unified 'model'.

For educational leaders, local governance and management within a public education system has required the integration and connection of their accountability to the local site and the system, the devolution of authority commensurate with their responsibility, and a clear understanding by both site leaders and state office of their mutual obligations and accountabilities.

Strengths of Partnerships 21

Participation in the policy shaping groups and the two surveys provided an opportunity for school and preschool leaders to record comments about the specific strengths and areas for improvement for the Partnerships 21 model.

In the survey at the end of 2000, 91% of schools and 87% of preschools registered general flexibility as the greatest strength of Partnerships 21, particularly in the deployment of staffing (48% of schools and 33 % of preschools) and the use of resources (58% of schools and 76% of preschools). The capacity to make local decisions to best suit the needs of children and students was also rated as a strength by a significant proportion of schools (71%) and preschools (65%). It was notable that 90% of the responses from schools and 84% of the responses from preschools sites did *not* identify an overall increase in resources as a major strength of Partnerships 21. The P21 leaders placed a significantly greater importance on the improved flexibility in using their sites' existing resources.

The global budget has allowed sites to use their financial resources more flexibly and many sites have found that they have been able to direct additional resources towards student learning through increases in teacher and school services officer time. In turn, this has seen the expansion of learning programs, in particular programs for improving literacy and numeracy, and to support students with special needs.

The model also reduced some of the restrictions on school and preschool staffing, and while teachers' industrial rights were not affected, schools have been granted greater freedom to select teachers and leadership staff. Schools have changed their leadership profile, increased time for specialist teachers and enriched curriculum offerings by appointing teachers to specific programs who best match the skills required for those programs.

Resource flexibility

Site leaders of Partnerships 21 schools and preschools have voiced few criticisms of the Partnerships 21 model itself. Instead their responses in the surveys emphasised a desire for a progressive increase in their local management of human, financial and physical resources. Their critique has mainly been about the need for greater support in the provision of timely and accurate information from state office, particularly reports on site expenditure on salaries, utilities and maintenance. A significant number of sites identified this need for greater information as being central to the full implementation of the Partnerships 21 model. (It is worth noting that the implementation of Partnerships 21 also coincided with the introduction by the Federal Government of a Goods and Services Tax (GST) creating additional workload for administrative staff.[2]) Requirements relating to establishing new procedures a the local site and developing the corresponding skills for local management also registered as an issue, but leaders also reported that the initial demands have abated as the changes have become embedded in schools' and preschools' routines and practices.

In a culture of local governance and management, the nature of the relationship between the head office and the local sites, alters in a number of significant ways. Essentially, local management redefines lines of responsibility and accountability for the local site and state office, whilst maintaining the operations of the department as a cohesive system.   Simultaneously, the department itself is required to re-engineer structures for the redistribution of resources – structures traditionally intended for the management of departmental ledgers. This has required a redefinition of accountability in ways that enhance the capacity of site leaders to manage additional resources whilst maintaining the required level of whole-of-department fiscal prudence and accountability

---

[2] A detailed analysis of the impact of the GST, Partnerships 21 and reforms to EDSAS is contained in the Office of Review report "Report on the review of the work of SSOs with administration and finance responsibilities".

At the operational level, the concerns by site leaders about the timeliness and accuracy of reports are indicators of their desire to make informed local decisions, and to provide credible and meaningful reports to their school communities. At the system level, local management has not reduced the accountability of head office to the government for the management of public money.

Further development of the processes for increased local flexibility with the need for effective financial management systems at the systemic level is currently underway. Most recently (March 2001) the vehicle for this reform process has been the Financial Management and Accountability Project, intended to clarify and redefine central budgeting, reporting and accountability processes in a context of increased site-based management of resources.

Quality Improvement

One of the key components of the Partnerships 21 model is quality improvement through strategic planning, with an emphasis on improving student achievement.

After twelve months of Partnerships 21, 78% of sites' responses sites to the survey at the end of 2000 indicated a moderate to significant level of change in quality improvement practices. Setting performance indicators and targets proved to be the aspect of strategic planning sites found most difficult, and they also registered the need for support with data collection and analysis. These aspects of planning were also highlighted as areas for development in quality improvement.

In South Australia there is growing evidence that school leaders and teachers are becoming more comfortable with using the learner achievement data, such as the "Basic Skills Test" data in reading, writing, spelling and numeracy, to reflect upon and improve their practices. This point was highlighted in the analysis of Partnerships Plans described in *Partnerships 21 – from plans into practice, the first six months* in August 2000. There remain however a number of questions about the use of data in analysis and improvement of the social outcomes of schooling, the place of locally developed analytical tools in school improvement, and ways of identifying the dimensions of schooling that provide the most 'leverage' in bringing about improvement in student learning.

In an echo of the issues surrounding increased local control of resources, local sites have greater authority to establish local policy, but do so within the context and obligations of systemic policies. They set their own strategic directions and performance indicators, but connect these to whole of system priorities, directions and accountability requirements.

Partnerships

The development of local governance is at the heart of the P21 model. The first year surveys indicated a modest level of change in schools and preschools practices in this area.

In the survey at the end of 2000, 36% of preschools and 59% of schools reported a significant increase in parent involvement in decision making. They also reported moderate to high levels of change in the governing council's role in strategic planning (preschools 54%, schools 54%) but limited change in their involvement in policy development or in developing partnerships with community organisations. 27% of schools reported that they were developing more effective governance structures, and 31% of schools a higher level of governing council input generally. Preschools, with a history of close community involvement in governance, registered lower rates of change overall.

In the context of the relatively short timespan, these findings bear out anecdotal evidence that policy development and establishing partnerships with other community groups has been less of a

priority for sites grappling with the more immediate challenges such as working through the resourcing and budgeting implications of entry into Partnerships 21.

Anecdotal evidence in 2000 also suggests that many schools sites, in particular, were deferring changes to their governing councils until the new amendments to the Education Act relating to the role of governing councils came into effect in 2001.

In the longer term, the success of the Partnerships 21 model will require the governing councils and principals and preschool directors to take shared responsibility for the governance of the school/preschool. This shared responsibility will involve the governing council in the determination and implementation of locally developed policy and strategic directions, and reporting to the local community on student outcomes and other performance indicators of quality. At the same time, the site leaders will be accountable to the Chief Executive for whole-of-system policies and priorities.

Looking forward

From the perspective of the leaders of Partnerships 21 sites, the benefits of the model clearly outweigh some of the early operational difficulties. However, our full realisation of the promise of Partnerships 21 will require principal/directors to have increased access to the management tools and data to strengthen their ability to make the connections between the key elements of the Partnerships 21 model. More sophisticated databases and more flexible reporting tools will enable them to better connect student learning outcomes with resource management and to better match their strategic directions with program outcomes and financial reports. This facility will also require tools that enable them to connect the various sub-components of their financial management responsibilities, in particular human resources, finances and assets.

As well, the future growth of Partnerships 21 as an integrated model of local governance and management will require on-going discussion and the development of enhanced understandings about the principal/director's role in integrating their accountability to the local community and to the system as a whole. This will occur as they strengthen the connections between the key elements of the model, partnerships, quality improvement, resource flexibility, in the movement towards improved student learning outcomes. In particular, this will involve the development of a sophisticated and dynamic partnership with the governing council – one that articulates the leader's direct leadership and management responsibly for student learning outcomes and financial management with the governing council's responsibility for approval, monitoring and reporting of the local site's policy, strategic plan and budget.

This is what we are looking forward to.

❖IEJ

Twin approaches to research: the Penn State and Flinders experience

Jonathan Anderson
School of Education, Flinders University *edja@flinders.edu.au*

Eunice N. Askov
Institute for the Study of Adult Literacy, Penn State University *ena1@psu.edu*

Two universities, The Pennsylvania State University in the United States and The Flinders University of South Australia, like many other higher education institutions, are developing online distance education programs. This paper reports on two approaches to offering a research methods topic at the postgraduate level, approaches that were determined in part by learners and their environment, which in turn determined modes of delivery. The modes of learning, development of learning resources, topic evaluation and feedback, and the lessons learned from offering the two research topics in an international context provide pointers to other higher education institutions planning to offer courses by flexible delivery.

flexible delivery, distance education, research methods, Internet, CD-ROM

AIMS AND PURPOSES

The Pennsylvania State University recently established a new campus, its 25th, which it designated the World Campus, to offer online courses to students anywhere in the world. As part of its program offerings, the first graduate degree program selected to go online was the Master of Education in Adult Education. Of the 33 credits required for the Master's program, students take 21 online, 3 credits being for a research methods topic called "Research Methods in Adult Education". This research methods topic, the first in the Master's program, was first offered in 2000.

At the end of 2000, Flinders University of South Australia entered into an agreement with the Rajabhat Institutes in Thailand to offer a Doctor of Education (EdD), a professional doctorate requiring three years of full-time study. The topics in the Flinders EdD are to be offered in a mixture of online and traditional teaching formats. The first online topic to be offered in 2001 is also a research methods topic called "Approaches to Research".

The purpose of this paper is to describe the context and delivery of these two research methods topics, one as part of the Penn State World Campus, and the other as part of the Flinders initiative in offshore teaching. In the case of Penn State, the research methods topic has already been taught, while in the case of Flinders University all preparations for teaching the research methods topic have been made. This paper compares the contrasting approaches with a view to drawing out lessons learned from flexible modes of course delivery in an international context.

LEARNERS AND THEIR ENVIRONMENT

Study at the graduate level in education is often undertaken on a part-time basis (Hammon & Albiston, 1998) while learners balance the demands of work, family, and study (Pitman 1997).

The challenge for higher education institutions is to adapt objectives, content and presentation for this group of motivated learners who have decided to return to study, usually to further their careers and to update their knowledge in educational methods and theories (Swenson, 1995).

Traditionally, university courses have been conducted by bringing students together in buildings and conducting instruction in laboratories and seminar rooms where lecturers or tutors try to guide learning activities to help students construct knowledge. According to Slavin (1997), this guiding process involves three key steps:

- understanding and knowing the prior knowledge of students so links can be made between old learning and new;

- responding to students to promote feedback and gathering the responses to use in the construction of new or refined knowledge; and

- helping students see how activities and ideas presented build or combine to increase understanding and knowledge.

Such an approach emphasises the active role learners take in knowledge construction. Learning thus defined has strong social aspects, with interactions between learners and tutors on the one hand, and between learners and learners on the other, playing an important role.

Where course delivery goes online, it is still important to maintain these kinds of social interactions since these have been shown to impact on learning outcomes (Burge, 1994; Jonassen et al., 1995). Computer based collaboration with its anytime any place capability is one means of achieving this (Harasim et al., 1995), while a key role for the lecturer is to act as facilitator in monitoring and supporting interactions (Wegerif, 1998).

The graduate students in the first Penn State World campus cohort were judged a typical adult education group. Many were returning to study, most were women, and all had other responsibilities. The students in the first Flinders entry to the offshore EdD program are all lecturers at Rajabhat Institutes (part of the university system in Thailand), pursuing doctoral studies to upgrade their teaching qualifications. For both groups of students, the online delivery of teaching programs allowed them to study while concurrently holding down full-time working positions.

METHODS OF DELIVERY

Penn State has been involved in outreach and distance education programs for more than a hundred years. The recent establishment of its 25th virtual campus, the World Campus, is a natural extension of its mission to serve the needs of distance students using new information and communication technologies for course delivery and to provide an interactive learning environment. The platform adopted for all its programs and courses is WebCT, which allows password protection together with such features as email, bulletin boards, chat rooms, and student record keeping.

Flinders University is a new entrant in the development of online distance education programs. The agreement with the Rajabhat Institutes in Thailand to offer a doctoral program for teaching staff led fairly naturally to a consideration of online delivery for course materials and instruction. The Approaches to Research topic, being the first topic offered in the program, is pioneering the School of Education's entry into distance education. WebCT was an option for developing the online version of the topic that thus far has been presented internally to students attending the university. For reasons that are elaborated in the next section, a decision was made to develop the

teaching and learning resources on CD-ROM and to link the CD-ROM to the School's Intranet for regular communication and interaction with students during the topic.

The choice of CD-ROM for the delivery of the main body of teaching and learning resources in the EdD program was made for several reasons. First and foremost, CD-ROM provides a portable resource allowing learning anytime, any place, at the user's pace, wherever there is access to a computer. In the rural regions of Thailand where the enrolled students reside, they are more likely to have access to computers with CD drives than access to the Internet. A second major consideration was cost for students. Once students are given the topic CD-ROM, there are no connection costs for accessing the learning materials.

The disadvantage of CD-ROM, of course, is that once a CD is pressed, the content is fixed. Nor does a CD allow any interaction of the kind described above between learner and tutor or between learner and learner. This disadvantage is overcome with the use of the Internet, which all students can freely access at the Rajabhat Institutes where they teach even if such access is not available on their home computers. The Internet allows email access to the tutor and fellow students, as well as access to a bulletin board developed especially for the program to cater for multiple threads and, indeed, multiple boards.

Because the CD-ROM is hot linked to the School Intranet, and since both CD-ROM and Intranet make use of browser software (Explorer or Navigator), the student, who is at a workstation linked to the Internet and with the course CD-ROM in the drive, can seamlessly navigate between the two sets of materials. Importantly, however, online costs are incurred only when using the bulletin boards or otherwise communicating with the tutor. Figure 1 shows the home screen for the CD-ROM produced for the topic, Approaches to Research. When students click to enter the Forum or view the Topic Schedule under Aims and Purposes (see Figure 2), they need an Internet connection. Otherwise, they are accessing the CD-ROM. Navigation between CD-ROM and Intranet is seamless.


Figure 1. Home Page for the topic - Approaches to Research

RESOURCES FOR LEARNING

Students enrolled at Penn State World Campus have access not only to academic tutors but also to a range of other learner support services. For technical support, for instance, often needed in the initial stages of a topic, there are dedicated staff to answer student queries. There is also access to Penn State's vast library resources.

A major problem encountered at the Rajabhat Institutes in Thailand, which for the most part are limited to offering bachelor degrees and a few professional master degree programs, are the severely limited library resources in each institution. The number of journal holdings is small and, of these, few are in English. Similarly the range of English textbooks for advanced studies in Education are virtually non-existent. This is not a problem peculiar to Thai universities but is common in many developing nations.
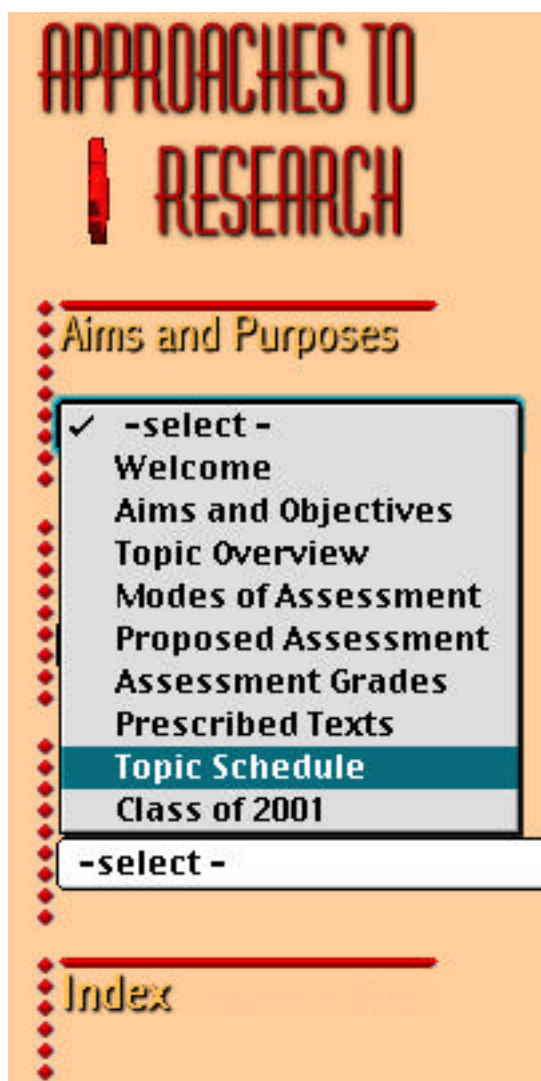
Figure 2. Pull-down menu options under Aims and Purposes

The very limited resources that Thai EdD students can call upon at their individual institutions poses a major difficulty for the offshore delivery of the Flinders EdD program. Nor was postage from Flinders University library a feasible option because of slow delivery times between Australia and Southeast Asia.

The choice of CD-ROM as a delivery option described in the previous section was seen as an ideal way of overcoming the problem of limited resources, achieved by building into the CD-ROM developed for the EdD research topic the beginning of a rich set of learning resources.

Figure 3 shows the range of resources from which students enrolled in Approaches to Research may select. "Glossary", for instance, takes students to a lexicon and concordance of terms used in statistics and empirical research in Education developed by a professorial colleague at Flinders University (Keeves, 1995). "Education Links" takes students to all higher education institutions in Australia, to universities worldwide, to major Australian education associations and organisations, and to major news media. Perhaps the most useful learning resource of all, particularly where library journal holdings are in scarce supply, is the link to "Educational Electronic Journals". Assembled here is an extensive list of more than 80 electronic journals from around the world that are scholarly, peer-reviewed, and accessible without cost. Penn State University has a link to this same list of educational electronic journals, which is placed also on the School of Education's web site (http://www.flinders.edu.au/education/). Other links under Learning Resources (see Figure 3) take students to "Research Proposals" presented by previous EdD students at Flinders University, to "Research Articles", "Reports and Monographs" published mainly by academic staff in the School of Education at Flinders University, to "Theses Online" completed recently in the School, and to an annotated bibliography of "Statistical Sites" containing useful information about statistical methods, statistical programs, and data sets.
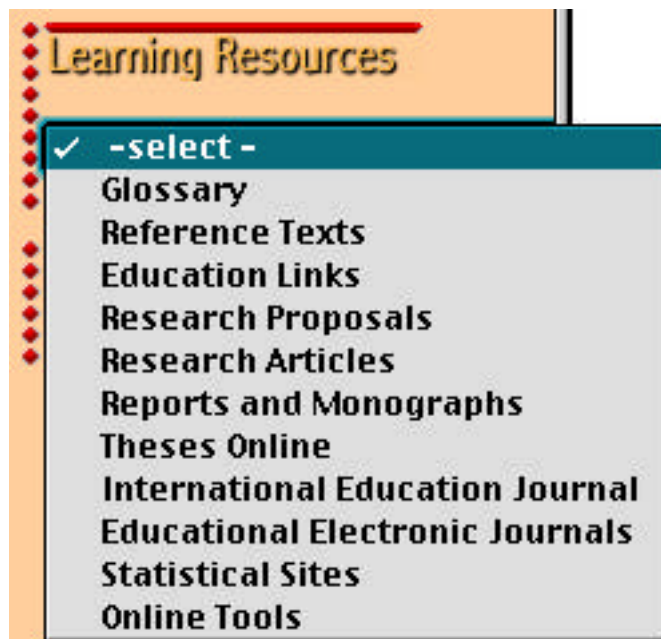
Figure 3. Pull-down menu options showing range of learning resources on CD-ROM

MODES OF LEARNING

Underlying both the Penn State and the Flinders University online delivery of research methods topics is a focus on creating a social learning environment where students can use other members of the group as a further learning resource. Hence for students at Penn State, the facility to use bulletin boards and email provided by the WebCT platform was an important element of the topic. Similarly for the Flinders program, a discussion forum and email are built in as core elements in teaching the topic. An additional advantage of the in-house discussion forum developed for Approaches to Research is the facility to email members of the group when new responses are added to any thread. Evidence suggests that this facility encourages students to log in to the forum more often since they know that further comments have been added.

As part of the Flinders offering, it is planned that two lecturers will travel to Thailand for a six-day face-to-face intensive teaching/learning session with all students enrolled in the topic. Since students come from different Rajabhat Institutes, most will also need to travel to the residential accommodation provided by the centrally located Nakhonsawan Rajabhat Institute. The major purpose of this residential teaching period is not to impart content since the CD-ROM will previously have been made available and textbooks will have been procured. Rather, the purpose is to develop close-knit interactive working relations among group members that may continue via email and the electronic discussion forum following the residential teaching period.

Penn State used primarily the main and group private bulletin boards as well as course email to achieve an interactive learning environment. While it experimented with chat rooms for the Master's research methods topic, these were used only occasionally by students in planning their group assignments; more often students preferred email to communicate and work together.

TOPIC EVALUATION

The first offering of the Penn State research methods topic was extensively evaluated (Askov, 2001b). A semantic differential type questionnaire was administered at the start and end of the topic. Among the findings, students reported that they felt more comfortable about the use of online teaching as a result of completing the topic. Another key finding is that students found group work useful in the sense that they felt less isolated and that they were participating members of a learning community with set goals and purposes.

Based on the Penn State experience, it is planned to evaluate the Flinders EdD topic taught offshore in Thailand using as a key dimension students' feelings of being at ease with the information and communication technologies in use, what Bandura and others call computer self-efficacy (Bandura, 1997; Eachus & Cassidy, 2001, Jablonski, 2001). An online version of the questionnaire to be administered has been prepared for administering at the start and end of the topic. Besides being used to evaluate the topic, the online instrument will serve as an example of one of the newer methods of research data collection, which is quite appropriate in a topic on research methods.

LESSONS LEARNED

The lessons learned from Penn State's World Campus delivery of the Master's degree research methods topic in adult education are valuable, not only for Flinders University as it embarks on the offshore delivery of its EdD program, but also for other higher education institutions planning to offer programs by flexible delivery. Askov (2001a) identifies institutional support as the most critical factor in delivering online programs and courses. Included in institutional support are technical services to deal with student queries, particularly in sorting out compatibility problems with equipment that often arise at the start of a topic. Institutional support is also critical in developing online course materials if faculty members are not to be unduly overloaded for engaging in distance education.

Askov (2001a) identifies social learning as another key factor if online programs and courses, especially at the graduate level, are to be successful.

Students who are geographically dispersed must be encouraged to form work groups to promote social learning. They must be encouraged to apply what they are learning to their own contexts and situations. Interactivity, on the part of the instructor with students and students with each other, is essential for learning to occur in distance education. (Askov, 2001a, p. 10)

What emerged from the Penn State experience, in particular, was that students said they felt they were part of an online community of scholars when they participated in World Campus courses.

As with all new topic development, it is essential to gauge student opinion about the effectiveness of innovations in teaching approaches and methods. At this relatively early stage in the use of the Internet, CD-ROM, and other communication technologies for teaching, there are many avenues for research if student learning and motivation are to be enhanced. Technological considerations ought not to override what educators know about student learning.

REFERENCES

Askov, E.N. 2001a. *Delivering the Adult Education Master's Degree Online through Penn State's World Campus.* Paper presented to the WebCT Asia Pacific Conference, Adelaide, South Australia, February 2001.

Askov, E.N. 2001b. *Researching Learning in Online Distance Education: A Study of Penn State's World Campus Adult Education Program*. Paper presented to the Flinders Institute of International Education, Flinders University, March 2001.

Bandura, A. 1997. *Self-efficacy: The Exercise of Control*. Freeman and Company, New York.

Burge, E.J. 1994. Learning in computer conferenced contexts: the learners' perspective. *Journal of Distance Education,* 9 (1), 19-43.

Eachus, P. and Cassidy, S. 2001. *The Computer Self-efficacy Web Site.* [Online]. Available: http://www.salford.ac.uk/healthSci/selfeff/selfeff.htm [2001, August 29]

Hammon, D.L. & Albiston, S.K. 1998. *Completing Graduate School Long Distance.* Sage, Thousand Oaks.

Harasim, L., Hiltz, S.R., Teles, L. & Turoff, M. 1995. *Learning Networks: A Field Guide to Teaching and Learning Online*. MIT Press, London.

Jablonski, D. 2001. *Analyzing Computer Self-efficacy of Teachers and its Effect on Curriculum at Hanoka's High School and Intermediate*. [Online]. Available: http://hale.pepperdine.edu/~djablons/665.casestudy.htm [2001, May 31]

Jonassen, D., Davidson, M., Collins, M., Campbell, J. & Bannan-Haag, B. 1995. Constructivism and computer mediated communication in distance education. *American Journal of Distance Education*, 9 (2), 7-26.

Keeves, J.P. 1995. *A Lexicon and Concordance of Terms Used in Statistics and Empirical Research in Education. School of Education*. Flinders University, Adelaide.

Pitman, V. 1997. *Surviving Graduate School Part Time*. Sage, Thousand Oaks.

Slavin, R.E. 1997. *Educational Psychology: Theory and Practice* (5th ed.). Allyn and Bacon, Needham Heights, MA.

Swenson, C.D. 1995. Graduate degree programs and distance education. In M. H. Rossman and M.E. Rossman (Eds) *Facilitating Distance Education: New Directions for Adult and Continuing Education*, No. 67. Jossey-Bass, San Fransisco.

Wegerif, R. 1998. The social dimension of asynchronous learning networks. *Journal of Asynchronous Learning Networks,* 2 (1), 34-49. [Online]. Available: http://www.aln.org/alnweb/journal/jaln_vol2issue1.htm#wegerif [2001, August 29]

A comparative study of family characteristics of Anglo-American and Asian-American high achievers

Philip O. Sijuwade
School Of Urban and Public Affairs, University of Texas, Arlington, Texas

The purpose of this study is to examine the similarities and differences in family characteristics of Asian-American and Anglo-American high achievers. Sixty matched parents were interviewed concerning four areas. The findings suggest that the family life of Anglo-American students tend to be less structured and provide less formal educational experience for children after school and on weekends. Similarities between these two groups include high parental expectation, concerned parents, stable family environment and close-knit family relationships.

Family characteristics, high achievers, Anglo-American, Asian-American, comparative study

INTRODUCTION

During the past two decades, the academic achievement of Asian-American children have been recognised not only be classroom teachers, school counsellors, and administrators, but also by the general public through the mass media. Successful stories of Indo-Chinese refugee children are frequently presented to the public as, for example, on such occasions as the Republican and Democratic National Conventions in 1984 and 1992 respectively. Their high achievement is reflected by their numbers as recipients of merit Scholarships, Westinghouse Science Talent Search Awards, and scholarships from leading universities, and has attracted the attention of many educators, sociologists who are focusing on the potential variables contributing to these accomplishments.

Facts regarding the academic commitment of these Asian-American students and their high scored in mathematics and science tests have been well documented. Asian-American students frequently scored higher than other students, particularly in the areas of mathematics and science (Fretchling et al, 1983; Hsia, 1991; Okada, 1984; and Peng et al, 1994). Several variables have been considered by researchers investigating this phenomenon:- attitude toward, school, parental expectation, cognitive development, gender difference, extra-curricular activities and parent-child relationships (Campbell et al, 1984; Hsia, 1991; Okada, 1984; Peng et al, 1994). Many educators and laymen have speculated that the accomplishments of Asian-American students are contributed to by their unique cultural traits in the above-mentioned areas and their traditional value system, which stresses the significance of education. Some characteristics of the learners, their parents, and their lifestyles have been examined by researchers (Peng et al, 1994). Yet the cause-effect relationship has not been investigated.

In order to determine whether the high achievements of Asian-American students are related to their cultural heritage, especially in the areas of family life, the researcher compared the family characteristics of high achievers of both Asian and Anglo-American students. The hypothesis of this study is that the process behaviours of Asian families producing high achievers are different from those of comparable Anglo families in some manipulable family variables.

LITERATURE REVIEWS

There is a body of research literature that addressed the topic of academic achievement and aits relationship to various factors, including motivation, family environment and intelligence quotient. As the United States becomes more culturally varied, questions arise as to whether these relationships are consistent across cultures. This particular aspect of academic achievement research is still in its infancy. Most of the studies focus on either black or white subjects, and their results and conclusions are confined to one ethnic group with little control on the variance of cultural force. thus, a cross-culture would be desirable to yield more valid information on the factors contributing to children's academic achievements.

Since influences at home are greater than those at school in effecting achievement (Conner, 1984, and Dolar, 1993), the researcher has narrowed the scope of study to the family characteristics of high achievers. From the reviewed literature, some familial elements are found to be influential on high achievers, especially in those of low-income black families.

The attitudes of parents toward education and the expectations of school teachers have made quite an impact on the learning results of their children (Gabel, et al, 1987). Feldman and Theiss' study (1992) reported that the more parents and students expected from school, the higher achievement the students attained. The more the parents expressed concern for their children's progress in school, the more the teacher matched that attitude (Goldman, 1983). Furthermore, the parents' belief in the use of schooling as enhancement for the child's future also contributed to his or her chances of success in school (Durkin, 1994). Parents often kept close contact with their children's teachers to insure the children's higher levels of achievement. The awareness of parents concerning the school program, no matter what their income, is conducive to high achievement (Scheinfeld, 1993; and Shade 1978). This parental contact with the school, however, is not always appreciated by children (Iverson, et al, 1991). Prior to the teenage years, the older the children are, the more they enjoy the extra parental attention. Once in high school, these children prefer to have more self-control and more autonomy concerning their school work.

Besides their expectation of school, the expectation parents have of their children also play an important role in children's achievement (Peng, et al, 1994). Such parental concern for achievement is reflected in pressure from the home and often leads to a close parent-child relationship (Shade, 1978). Shade's study (1978) found that the interaction between parents and their children has more to do with the child's success than does the socio-economic level of the family. the parents' expressions of warmth, support, interest, affection, and encouragement were found to have a great impact on high achievers from low-income black families (Dolan, 1993). The assistance of parents in establishing performance goals and in giving guidance with perceived problems is another trait common to black high-achieving students with a low socio-economic background (Shade, 1978).

Parents' expectations of education, teachers and children, and their relationships with their children, directly or indirectly often affect the type of extra-curricular activities a child engages in after school (Shade, 1978). Peng and others (1994) have found that Asian students were more likely to participate in honorary or subject matter clubs than in vocational educational clubs, church activities or athletics. in addition, Asian children tend to make friends with people who have similar racial and cultural backgrounds, a tendency that becomes stronger with age (Hsia, 1991). the type of after-school activities a child engages in and the friends a child associates with naturally have a considerable degree of influence on academic performance. For example, Asia subjects in Peng's study were likely to out-perform their white counterparts. the study of Riley-Johnson and others (1993) indicated that children who watched less television tended to have

better grades in school and developed higher IQ's thus parents could contribute to their children's academic success by limiting time spent watching television.

The effects of time spent on homework after school, like television watching, is another controversial issue debated by many educators. Keith (1992) reports, "An increase in time spent on homework has a positive effect on a student's grades in high school." By studying successful black readers in urban schools, Durkin (1994) has found that these children had help from family members, usually mothers and older siblings.

The maintenance of some structure and order for children is conducive to their school performance (Shade, 1978). Regularity in daily life and family stability can be considered part of this variable. However, divorce is not found to be as detrimental to children's achievement as previously thought. Hommand (1989) reports that divorce only affects boys' mathematics scores. For immigrants, the length of time living in the United States also affected their verbal skills and scores on all tests. Peng et al (1994), found that children are usually able to overcome the language barrier after residing in the United States or six or more years.

Other variables such as parents' education and number of siblings might also affect children's achievement in school. However, no conclusive results have been reported.

METHOD

In order to examine possible cultural factors, this study set out to compare the family characteristics of both Asian and Anglo high achievers whose scores on achievement tests were over the ninetieth percentile.

Sample

In order to control the socio-economic factor, all the subjects were selected from three schools located in a middle and upper class suburban neighbourhood. In this particular school district in north-east Texas, over 10 per cent of the student body is Asian-American. Stratified random sampling was employed and 30 students were first selected from the Asian-American population in the ninetieth percentile or above on their *Texas Assessment of Academic Skills*. Anglo students were then chosen to match with the Asian sample according to sex, grade level and scores on achievement tests. All 60 students were selected from Grades 5 through 11 with the average grade level of 7.4. The students consisted of 32 girls and 28 boys. The Asian-American sample was made up of students of Chinese, Korean, Filipino, Vietnamese and East Indian extraction. Sixty per cent of the mothers of those in the sample were employed outside the home at the time of the interview. Half of the Asian-American subjects were foreign born with an average of 5.3 years on arrival in the United States. Two-thirds of the sample had only one sibling. Coincidentally, six Asian families and six Anglo families were headed by a single mother.

Procedure

The data were collected through structured interviews of selected children's parents held at either the sample's home or the researcher's office as they preferred. in the Anglo sample, two-thirds of the interviews, 17 were with mothers only, seven were with fathers only and six were with both parents.

All the interviews were conducted without the children. The researcher was always well-received, and parents were very co-operative in answering each question in detail. The interviews lasted an average of 90 minutes. Some of the parents were so concerned about education that they took advantage of this opportunity to discuss the educational needs and problems of their offspring.

Instrumentation

In addition to the demographic data form, the researcher used an open-ended questionnaire for the interview. This questionnaire consisted of 17 items related to (a) parents' attitude toward education, their expectation of school, teachers and their children; (b) parents' contact with schools and interaction with their children; (c) children's after-school activities; and (d) lifestyle of the family.

Limitation

Since the subjects were selected from a homogenous middle and upper-class environment, a majority (60%) of the families were headed by well-educated parents with a double income. It is possible that the results of this study are affected by the socio-economic status of the interviewed parents. Since parents participated voluntarily, those who declined to be interviewed may have been less concerned with their children's education. Such lack of voluntary participation if often seen in social science research studies and is unavoidable.

RESULTS AND DISCUSSION

The results of this study are presented according to the areas covered by the reviewed literature and the 17-item questionnaire.

Parents' Attitudes Toward Education and Expectations of Those Involved

All the parents interviewed express a deep interest in their children's school and valued education highly. Six Anglo parents even indicate their support of the "back to the basics" movement in Texas. Asian parents are less vocal about the educational goals and policies in the United States. In general they are pleased with what schools have offered their children and what teachers have taught in the classroom. All of them indicate that their children are interested in school and learning. Such comments are voiced by 80 per cent of the Anglo parents. Twenty per cent of the Anglo parents complain that their children become bored in class due to poor teaching techniques and unchallenging learning materials. Such discrepancy can be the result of the different cultural background; Asian parents tend to respect teachers more than the Anglo parents do and avoid a challenge to the teacher's authority. However, Asian parents show stronger feeling than Anglo parents concerning the value of education for their children's success because they were willing to commit all their resources to insure the best education available for their offspring.

Different degrees of parental expectation for their children are found between Asian and Anglo-American parents. All Asian parents report that they expect their children to make an average grade of "A". only two-thirds of the Anglo-American parents expect straight "As" from their children, while the remaining one-third are willing to accept "Bs" as the minimum grade. Four Anglo-American students reportedly are not serious about their report cards. All the Anglo parents are pleased with their children's performance in school. while half of the Asian parents feel that their children do not perform well all the time. This might demonstrate that Asian parents are more demanding and have a higher expectation for their children. It is possible that this attitude relates to the educational level of these parents; all the Asian parents have college degrees as compared with 56 per cent of the Anglo parents.

In the areas of career choice, some variations are found between Asian and Anglo parents. nearly half of the Asian parents (46%) hope their children would choose the medical field compared with 29% of the Anglo's. It is possible that due to their experience as immigrants the Asian parents realise that jobs in the medical field would be more secure and would encounter less interpersonal conflicts. When the parents were asked whether they had attempted to direct their children toward

certain occupations, only two Asian parents and 12 Anglo parents said, "No". It would appear that most Asian parents still preserve the traditional attitude that parents would play a major role in their children's education and career choice (Stigler, et al, 1992). therefore, they try very hard to assert their opinions and to exert an influence on their children's future occupation. On the contrary, individualism is highly marked and valued by Anglo parents and personal choice is generally well respected.

Parents' Interaction with School and Children

Both Asian and Anglo parents are well informed and knowledgeable about what is going on in school, such as school work, projects, social activities, and various academic competitions. this is a major factor in children's success in school (Dolan, 1993). Yet Anglo parents make more school visitations and contracts with teachers than Asian parents. this is a very desirable phenomenon as the study of Gabel and others point out (1987). It seemed that Anglo parents are more active in school functions than Asian parents who assume a supporting role. the busy working schedule of both immigrant parents, the language barrier and a reserved of shy cultural nature all contribute to these behavioural patterns.

Nevertheless, these minor differences seem to have little impact on the relationship between parents and their children. Half of the Asian parents indicate that they are always involved in their children's homework and projects while only one-third of the Anglo's are so involved. Fifty-seven percent of the Anglo parents express the opinion that their children are either very independent or are only given assistance as needed. Only a little over one-third (36%) of the Asian parents consider that their children belong in this category. Asian parents traditionally feel obligated to assist their children in any way they can (Stigler, et al, 1992). This parental commitment carries over into non-academic activities as well.

When the selected parents were asked "In what area(s) do you feel that your child is more influenced by peers than by you?", 45 per cent of the Asian parents claim none. Thus, thanks to close parent-child "team-work", these parents still have an influence upon their children's social development and maintain more control than do the child's peers in the areas of clothes, courses of study, food preference and extra-curricular activities. Eighty-six per cent of the Anglo parents reveal that their children are more influenced by their peers in the area of clothing. It is interesting to note that only 50 per cent of the Asia parents give the same response. In general, Asian parents seem to maintain more control over their children than Anglo parents do. Such a behavioural pattern contradicts the findings of Scheinfeld (1993). His research indicate s that low-income, black low achievers have more maternal constraints. it is very possible that parental SES plays a part in such differences.

The parent-child relationship is also disclosed when the parents describe their daily routine activities. Ninety per cent of the interviewed families have supper together, except when there is a scheduled conflict involving sports practice or music lessons. In two families from East India the father regularly eats separately. No other significant differences are found between the Asian and Anglo families with respect to daily activities. in comparing weekend activities, it is found that Asian families tend to be more regular, rigid, and task-orientated than the Anglo families. Cultural activities, private music lessons and language school are mentioned more by Asian parents as a major portion of their children's weekend activities. Anglo parents tend to provide children with a larger variety of experiences which are not only flexible but also fun, such as shopping, going to movies, boat repair, and so on. However, regardless of the specific nature of weekend activities, it is found that both Asian and Anglo, at least one of the mentioned activities involve the entire family.

Children's After-School Activities

The extra-curricular activities listed by the interviewed parents are different for the Asian and Anglo children. Music lessons rank first among Asian sample followed by sports. Sports are first for the Anglo children followed by church activities. it appears that Asian children follow a more individualised pursuit while Anglo children lean towards group-orientated venture, such as drill team, debate and swimming.

Confucianism greatly influences the Asian culture, since it values the mind more than the body. efforts to refine an individual's mind have won greater attention. Some intellectual pursuits of Asian children are related to their ethnic heritage. The Asian children's individualised extra-curricular experience may have a certain effect on interaction with friends after school. if the frequency of phone calls from or visits by peers is any indication, the Asian child has less social contact than the Anglo child. An explanation for this might be the busy after-school schedule which limits the social life of the Asian child. The cause-effect of Asian children's social skills is not clearly identified, yet it is often observed by teachers (Campbell, et al, 1984).

All but one of the parents consider homework the major after-school activity, with the number of hours spent varying according to the age of the child. Some of the children start their homework right after school while others wait until evening. Since homework is given priority, television watching is limited to an average of one hour per day for all children, much the same for Asian or Anglo. this result supports the reviewed research on homework and television watching (Keith, 1992).

On weekends, the activities engaged in by Asian and Anglo students include all cognitive, affective and psychomotor domains with a different degree of emphasis in each domain; Asians tend to engage in activities related to cognitive learning while Anglos are more interested in affective and psychomotor development. Although all children are expected to help with household chores, Asian children are usually assigned less responsibilities than their Anglo counterparts. this difference reflects the cultural characteristics of both groups; since disciplining the mind is the top priority for educated Asians, one's labour is considered less worthy. In fact, Asian parents tend to avoid bothering children with household chores and would rather they spend more time in study (Stigler, 1992).

Lifestyle

Regardless of parents' marital status, most of the families are very stable; 30 per cent of the families had not moved while ten per cent had moved twice since the child began school. It is interesting to note that eight out of 12 single mothers do not intend to remarry during the child-rearing process; they want to avoid any more major changes in the children's lifestyle. Actually, Fowler and Richards (1988) pointed out that the father's absence did not have a great impact on children's academic achievement.

Children benefited from a regular lifestyle with a feeling of stability and sense of control. similar results have been found in studies on black high achievers (Durkin, 1994). the children studied tended to set goals and plan for their future. Even on weekends, most of them (63%) vary their bedtime by only one hour. Their regular lifestyle is consistent concerning meal times as well. Most of the children (95%) have breakfast every morning; the exceptions are two Anglo girls; one wants to control weight and the other often left no time for breakfast after fixing her hair. Generally speaking, no significant difference was found between the Asian and Anglo families in this area.

CONCLUSION

Although minor differences and variation in parenting style are found in child rearing, basically both Asian and Anglo parents in the United States share a similar value system; actually there are more similarities than differences. Both groups are concerned, warm, interested and supportive parents and work to maintain a close relationship with their offspring. Divergence in parental expectation and control seems to have little impact on children's achievements. it is unclear whether the parental expectations asserted by the Asian parents have actually promoted their children's achievement. in other words, it is questionable whether those identified Asian high achievers could have reached the academic excellence without parental push and assistance. According to several studies, children's grades can be improved with parents' support and assistance (Comer, 1984; Gabel, 1987; Goldman, 1983). Perhaps this explains why the selected Asian children are over-achievers. A stable lifestyle might be another major contributing factor to high achievement ion school. In this study, parents have demonstrated their concern and support by making a special effort to maintain an even lifestyle for their children.

REFERENCES

CAMPBELL, J.D. et al, 1984 Impact of Ethnicity on Math and Science among the gifted. Paper presented at the annual meetings of the *American Educational Research Association*, New Orleans, La.

COMER, J. 1984 Home School relationships as they affect the academic success of children. *Education and Urban Society*, 16(3), 323-337.

DOLAN, L. 1993 The prediction of reading achievement and self-esteem from an index of home educational environment: a study of urban elementary students. *Measurement and Evaluation in Guidance*, 16(2), 86-94.

DURKIN, D. 1994 Poor black children who are successful readers. *Urban Education*, 19(1), 71-81.

FRETCHLING, J.A. et al, 1983 Summary of MCPS data on difference in minority performance and participation. Rockville, Maryland: Montgomery County Public Schools.

FELDMAN, S. 1992 The teacher and students as pygmalions: joint effects of teachers and student expectations. *Journal of Educational Psychology*, 74(2), 217-223.

FOWLER, P.C. and RICHARDS, H.C. 1988 Father absence, educational preparedness, and academic achievement: a test of the confluence model. *Journal of Educational Psychology*, 70(4), 595-601.

HABEL, H. 1987 Parent-teacher communications in relation to child academic achievement and self-concept. *Peabody Journal of Education*, 15(3), 142-145.

GOLDMAN, R. 1983 Cross-cultural adaption of a program to involve parents in their children's learning. *Child Welfare*, 52(8), 521-532.

HSIA, J. 1991 *Cognitive Assessment of Asian-Americans*. Princeton, N.J.; Educational Testing Service.

HAMMOND, J.M. 1989 Children of divorce: a study of self-concept, academic achievement, and attitudes. *The Elementary School Journal*, 80(2), 55-62.

IVERSON, B.K. 1991 Parent-teacher contacts and student learning. *Journal of Educational Research*, 74, 394-396.

KEITH, T.Z. 1992 Time spent on homework and high school grades: a large sample path analysis. *Journal of Educational Psychology*, 74(2), 245-253.

OKADA, T. 1984 *SAT Primer; 1982 test results*. Washington, D.C.: Office of Planning, Budget and Evaluation, U.S. Dept. of Office of Planning, Budget and Evaluation, U.S. Department of Education.

PENG, S. E. et al, 1994 School experiences and performance of Asian-American high school students. Paper presented at the annual meeting of American Educational Research Association, New Orleans, La.

RIDLEY JOHNSON et al, 1993 The relation of children's television viewing to school achievement and I.Q. *Journal of Educational Research*, 76, 294-297.

SCHEINFELD, D.R. 1993 Family relationships and school achievements among boys of lower-income urban black families. *American Journal of Orthopsychology*, 53(1), 127-143.

SHADE, B.J. 1978 Social-psychological characteristics of achieving black children. *The Negro Educational Review*, 29(2), 80-86.

STIGLER, J.W. et al, 1992 Curriculum and achievement in mathematics: a study of elementary school children in Japan, Taiwan, and the United States. *Journal of Educational Psychology*, 74(3), 315-322.

The multilevel analysis of students' achievement in learning the Chinese language

Ruilan Yuan
Flinders University of South Australia *Rlan.Yuan@flinders.edu.au*

John P. Keeves
Flinders University of South Australia *john.keeves@flinders.edu.au*

A three-level hierarchical or multilevel model is employed to examine the factors that influence students' achievement in learning the Chinese language. The factors might be many or various, such as school factors, factors related to teachers, parents and peers. This study, however, only examines student level factors. The first level of the three-level multilevel model is the within-student level; the second level is between-student level; and the third level is between- class level. The results show that different factors at different levels of the model have effects on the students' achievement. Interaction effects are also observed between some variables and the achievement.

Multilevel Analysis, Student Achievement, Chinese Language

INTRODUCTION

After the World War II, especially since the middle of 1960, when Australia's involvement in business affairs with some Asian countries in the Asian region started to occur, more and more Australian school students started to learn Asian languages. The Chinese language is one of the four major Asian languages taught in Australian schools. The other three Asian languages are Indonesian, Japanese and Korean. In the last 30 years, like other school subjects, some of the students who learned the Chinese language in schools achieved high scores in learning the Chinese language, and others were poor achievers. Some students continued learning the language to Year 12, while most dropped out at different grade levels. Therefore, it is considered worth investigating what factors influence student achievement in the Chinese language. The factors might be many or various, such as school factors, factors related to teachers, classes and peers. This study, however, only examines student-level factors that influence achievement in learning the Chinese language.

DESIGN OF THE STUDY

The subjects for this study were 945 students who learned the Chinese language as a school subject in a private college of South Australia in 1999. The instruments employed for data collection were student background questionnaire and attitude questionnaire, four Chinese language tests, and three English word knowledge tests. All the data were collected during the period of one full school year in 1999.

THE THREE-LEVEL HLM ANALYSIS

It should be noted that one of the aims of this study is to examine the factors influencing students' achievement in the Chinese language across school grades and over time. The factors to be examined are categorized into three types: (a) *intra-student* factors; (b) *inter-student* factors; and

(c) *environmental* factors of the school class group. The single-level PLS analysis is clearly insufficient to identify differences in the factors between students within year level groups and among the year level groups of students. Therefore, a three-level multilevel model is employed to undertake this task.
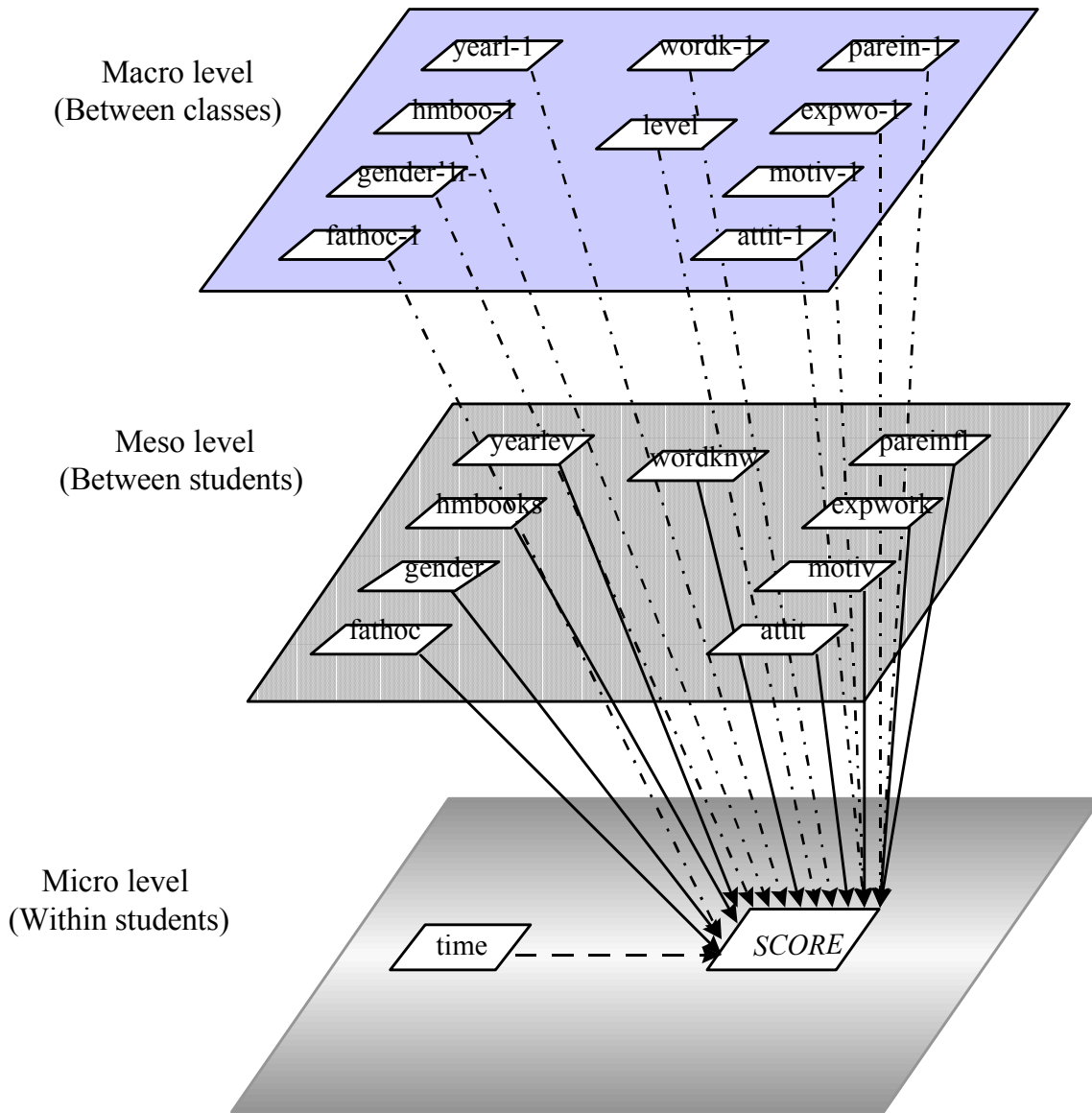
Bryk and Raudenbush (1992) have identified three general purposes of hierarchical or multilevel model applications: (a) to improve the estimation of effects for individual units; (b) to formulate and test hypotheses about across-level effects; and (c) to partition variance and covariance components among levels. The analyses undertaken in this study using a three-level multilevel linear model involves the investigation of student, as well as class factors influencing students who are learning the Chinese language. The data for the analyses are organized at the within- and between-student levels as well as at the between-class level. It is necessary to note that three different terms are used in the analyses and interpretation of the data. The first level of the three-level multilevel model is referred to as the micro level, or the within-student level, or Level 1; the second level is referred to as the meso level, or between-student level, or Level 2; and the third level is referred to as the macro level, or class level, or Level 3. Hence, the present paper discusses the results obtained from the analysis of the three-level multilevel model to show whether or not differences can be observed with respect to the within- student factors, the between-student factors and the between-class factors that are hypothesized to influence student achievement in learning the Chinese language. The outcome variable is *SCORE* (see Figure 1).

Raudenbush and Bryk (1994) have stated that hierarchical linear models (HLM) enable the testing of hypotheses about effects that occur within and between each level of the variables and about the interrelations between them. In other words, hierarchical linear models allow for the analyses of data at different levels simultaneously.

Figure 1 presents the hypothesized three-level hierarchical or multilevel model of student level factors influencing the achievement in learning the Chinese language across year levels over time. Lietz (1995, p. 121) argued, "It is not possible to form latent variables in the actual HLM analysis and manifest variables are employed instead". Hence, in the current hypothesized three-level multilevel model, some manifest or observed variables (see Table 1) were selected as variables for the HLM analyses.

The first level is the *intra-student* model, namely the within-student model (micro level model). *Time* is used as a predictor to estimate each individual student's four scores and to explore the differences within each student at the four time points. The second level is the *inter-student* level model, that is the between-student model (meso level model) which is used to compare the differences in the achievement between students averaged over the four occasions. Level three is the macro level model or the between-class model, which allows for the effects of the average levels of student characteristics between classes. Therefore, the variables at the class level are aggregated from the between-student level, namely the second or meso level, to address the effects of differences in the composition of the class environment.

Figure 1 combines the intra-student, inter-student and environmental factors that are hypothesized to influence student achievement in learning the Chinese language. In the model, it is hypothesized that, at the within-student or micro level, the variable hypothesized to influence the achievement is referred to as time or occasion. At the between-student or meso level or Level 2, all the nine variables are hypothesized to influence directly student achievement in learning the Chinese language (cachment), namely, fathoc, gender, hmbooks, yearlev, wordknw, pareinfl, expwork, motiv, and attit.

Figure 1   A Hypothesized Three-level Multilevel Model of Factors Influencing Student Achievement in Learning the Chinese Language

The variables at the between-class or macro level are averaged and aggregated from the second level, namely the between-student level. In addition to the nine variables aggregated from the second level, one more variable is created and used at the between-class level, namely, grade level (level).

At the micro level or Level 1, timeid is used as a predictor to examine the score differences within each student at the four time points. At the meso level, namely Level 2, the differences in the achievement between students on the four occasions are compared. Nine latent variables are hypothesized to influence the outcome variable score. These nine variables are fathoc, gender, hmbooks, yearlev, wordknw, pareinfl, expwork, motiv and attit. At the macro level or Level 3, the comparisons are made between the average levels of student achievement among classes. Ten variables are proposed to influence score, among which nine variables are aggregated and averaged from the variables at the meso level or Level 2 (i.e. fathoc-1, gender-1, hmbook-1, yearle-1, wordk-1, parein-1, expwo-1, motiv-1, and attit-1) with one additional variable (i.e. level) added to the model so that the effects of grade or year levels can be examined. In other words, with the class level model, it is also possible to investigate whether or not differences exist between the grade or year levels with respect to their achievement in learning the Chinese language. It should be noted that this analysis does not consider the relative performance of each individual class group, or the effects of the teachers on the performance of each class group.

Several preliminary steps of analysis are required prior to the conduct of the main analysis of the three level hierarchical or multilevel model. First, variables must be selected from the results of the PLS analyses to be included in the regression equations. Second, the data for analysis need to be prepared in an appropriate form for analysis at the three levels, and finally the models must be specified. The following sections describe these three steps respectively.

Variable Selection

HLM analysis employed the observed or manifest variables. Therefore, all the 20 manifest variables that form the nine latent variables are selected as the variables in the three level HLM analysis. Consequently, the variables used in the equations in the fully unconditional model and the final model of the HLM analysis are the corresponding manifest variables of the latent variables used in the hypothesized three level HLM model.

It is important to point out that 12 items from the Student Background Questionnaire, and five items, namely *chiease* (ease of learning the Chinese language)*, chiefft* (effort in learning the Chinese language)*, intrest* (interest in learning the Chinese language)*, chiuse* (usefulness of learning the Chinese language), and *schefft* (effort in schoolwork), which indicate the five attitude sub-scale scores obtained on both Occasions 1 and 2 are included in the analyses. The criterion variable in the three level HLM model is *score* which indicates the four separate scores from the Chinese language achievement tests. Table 1 provides the information on all the variables employed in the three level HLM model.

Data Preparation

In order to separate out the variables for Level 1 or the micro level, a new file was built using SPSS for Windows Version 6.13 (Statistical Package for the Social Sciences, 1993). The first variable was *timeid* indicating Terms 1 to 4 for each individual student. The second was *classid* representing Classes 1 to 39. The third variable was *studid* representing each individual student. The final variable was the *score* for each of the four occasions which was the scaled Rasch scores for the four terms for each individual student. Hence, in general, each student had four Rasch scaled scores, with initially 945 students. However, four Year 4 classes (N=121) and 27 Year 12

students only completed two of the four tests (see Chapter 5). The scores for the other two terms were treated as missing data.

Table 1   The Variables Included in the HLM Model

Level 1 Variable (Intra-student Level)	Coding
▪ *timeid* (The 4 term points in the 1999 school year)	Combined scale
Level 2 Variable (Inter-student Level)	
Fathoc ▪ *focc* (Father's occupation)	1 – Not in work force, 2 – Unskilled, 3 – Semiskilled, 4 – Skilled workers, 5 – Clerical workers, 6 – Managerial, 7 – Upper and lower professional.
Gender ▪ *sex*	1-Male; 2-Female
Yearlev ▪ *yearlevl* (Grade level)	Year 4 – Year 12
▪ *yearchin* (No. of years learning the Chinese language)	1 – Just started, 2 – 1 to 2 years, 3 – 3 to 4 years, 4 – 5 to 6 years, 5 – 7 or more years.
Hmbooks • *books* (Books in general at home)	1 – 0-10, 2 – 11-25, 3 – 26-100, 4 – 101-200, 5 – more than 200.
▪ *chibooks* (Chinese language books at home)	1 – None, 2 – 1 to 5, 3 – 6 to 10, 4 – 11 –15, 5 – 16 or more.
Pareninfl ▪ *encourag* (Encouragement from parents)	1 – Hardly ever or never, 2 – Occasionally, 3 – Often.
▪ *parfeel* (Parents' feeling about their child's learning the Chinese language.)	1 – I don't know about their feelings. 2 – They are not in favour of it. 3 – They don't care about it. 4 – They are in favour of it.
Wordknw • *scorknw* (The Rasch scores obtained from English Word Knowledge tests.)	The average Rasch score on English word knowledge tests.
Expwork ▪ *expectoc* (Students' expected occupation)	1 – Not in work force, 2 – Unskilled, 3 – Semiskilled, 4 – Skilled workers, 5 – Clerical workers, 6 – Managerial, 7 – Upper and lower professional.
Motiv ▪ *schefft1, schefft2* (Effort into schoolwork on Occasions 1 & 2)	Combined scale
▪ *homchin* (The Chinese language homework)	1 – I do not do any Chinese home-work, 2 – About one hour, 3 – About two hours, 4 – About three hours, 5 – Four or more hours.
▪ *homewrk* (Homework for all the subjects)	1 – I do not do home- work, 2 – Less than 2 hours, 3 – Less than 5 hours, 4 –5 to less than 10 hours, 5 –10 to less than 15 hours, 6 – 15 or more hours.
Attit ▪ *chiease1, chiease2* (Ease of learning Chinese Occasions 1 & 2)	
• *chiefft1, chiefft2* (Effort in learning Chinese Occasions 1 & 2)	Rasch scaled scores
• *intrest1, intrest2* (Interest in learning Chinese on Occasions 1 and 2)	
• *chiuse1, chiuse2* (Usefulness of learning Chinese on Occasions 1 and 2)	
Level 3 Variable (Class level)	
▪ *level-1* (Grade or year level)	1 to 8 grade or year level
▪ *and all Level 2 variables aggregated*	
Outcome Variable	
▪ *score*	The Rasch scaled scores for Terms 1, 2, 3, and 4 respectively.

Likewise, where data were missing for individual students or other class groups, the missing data code was assigned. As a result of these missing data, the four Year 4 classes and 27 Year 12 students had only two scores available for analysis but could be included in the analysis. However, another 12 students had inadequate data with scores for only one occasion, and were therefore deleted from the analysis. The scores available for the HLM analyses involved 3,109 cases at Level 1, 933 students at Level 2, and 39 classes in eight grade levels at Level 3. The Rasch scores at Level 1 formed the outcome variable (*score*).A second file was also created in SPSS for the between-student data at the meso level. Since students differed in the years that they had been learning the Chinese language, a variable *level* was formed that indicated years of learning Chinese. At Level 2 or the between-student level, all the manifest variables forming the nine latent variables used in the PLS analysis (i.e. *fathoc, gender, hmbooks, yearlev, wordknw, pareinfl, expwork, motiv* and *attit*) were also selected together with a new variable *level* for the HLM analysis. In addition, each student had a student -id and a class-id to identify the student and the class to which the student belonged. In should be noted that the attitude scores obtained from *chiease, chiefft, intrest*, and *chiuse* on Occasion 2 were also used in the HLM analysis.

At Level-3, or the between-class level, all the ten variables at Level-2 were aggregated and averaged to become Level-3 variables (see Figure 3) by running the command 'Aggregate' in the SPSS program. This was achieved by using the class identification number as a break variable. Moreover, there was a variable that indicated the class level at which the Chinese language was being studied.

In addition to the above steps, the following procedures were also taken to prepare data for three-level HLM analysis. First, all the files created were saved in SPSS as *INTRASTU*.sav for the Level 1 file, *INTERSTU*.sav for the Level 2 file, and *CLASSLEV*.sav for the Level 3 file. Secondly, *TIMEID* was used as the micro level identification variable, *STUDENTID* was used as the Level 2 identification variable, and *CLASSID* was used as the Level 3 identification variable. Finally, a sufficient statistics matrix (SSM) was created and used as the data file for the three-level HLM analysis of the Chinese language achievement the data. A '*Pair-wise*' procedure was used at Level 1 for the calculation of the SSM because it was identified as containing missing data.

Missing Data

It is worth noting that the HLM program accepts missing data at Level 1, namely the within-student level. The missing data were replaced with the grand mean at Level 2, that is at the between-student level. Consequently, the aggregated data at Level 3 had no missing data.

Model Specification

The analysis of a three-level multilevel model involves different steps. In general, the first step is to run a fully unconditional model in which no predictors are specified at either Level 1 or 2, nor at Level 3. The second step is to add into the equation the micro level or Level 1 variables. The third step is to add the meso level or Level 2 variables that are statistically significant. The fourth step is to add the macro level or Level 3 variables that are significant, and run a final model in which micro, meso and macro level variables are all entered into the equation. The sections that follow discuss the fully unconditional model and the final model.

The fully unconditional model

Bryk and Raudenbush (1992) have stated that the fully unconditional model is the simplest model among all multilevel models and contains no predictor variables from any level. The fully unconditional model is used to obtain the estimates of amount of variance explained at each level in the model. In this study the fully unconditional model for the three levels is written as

Equations 1, 2 and 3. At Level 1, the achievement score in the Chinese language for each student is modelled as a function of the expected achievement plus a random error. The model is represented as:

$$Y = P_0 + E \qquad \text{[Equation 1]}$$

where

Y = the Rasch test score for achievement in the Chinese language as outcome variable;

P_0 = the intercept;

E = a random within-student effect which is assumed to be randomly distributed with a mean of zero and a variance of σ^2.

The model at the meso level or Level 2 is given by equation:

$$P_0 = B_{00} + R_0 \qquad \text{[Equation 2]}$$

where

P_0 = the intercept;

B_{00} = the mean Rasch test score for achievement at Level 2;

R_0 = a random between-student effect which is assumed to be randomly distributed with a mean of zero and a variance of τ_π.

The macro level or Level 3 model is specified as:

$$B_{00} = G_{000} + U_{00} \qquad \text{[Equation 3]}$$

where

B_{00} = the mean Rasch test score for achievement at Level 3;

G_{000} = the grand mean Rasch score across classes;

U_{00} = a random effect associated with classes which is assumed to be randomly distributed with a mean of zero and a variance of τ_β.

The relevant results of the variance components generated by the fully unconditional model for the three-level HLM analysis are given in Table 2.

The fully unconditional model makes it possible to partition the total variability in the outcome measure Y into components at these levels, namely the within-students, σ^2; the between-students, τ_π; and the between-classes, τ_β. In addition, it also allows for estimations to be made of the proportions of variance analysed at each of the three levels. The following three equations were adapted from Bryk and Raudenbush (1992, p. 177).

Table 2   Final Estimation of Levels 1, 2, and 3 Variance Components in the Fully Unconditional HLM Model

Random Effect		Standard Deviation	Variance Components	df	Chi-square	P-value
Level-1,	E	1.303	1.698			
INTRCPT1,	R_0	0.931	0.867	893	1359.95	0.000
INTRCPT1/INTRCPT2	U_{00}	0.992	0.984	38	588.79	0.000

Notes: $E = \sigma^2$ $R_0 = \tau_\pi$ $U_{00} = \tau_\beta$

Deviance = 11489.25 with 4 parameters

At Level 1, or micro level, the proportion of the variance in the Chinese language achievement that exists between students in classes and within students is given by:

$$\frac{\sigma^2}{\sigma^2 + \tau_\pi + \tau_\beta} = \frac{1.698}{1.698 + 0.867 + 0.984} = 0.479 \ (n=3,109)$$

At Level 2, or meso level, the proportion of the variance in the Chinese language achievement that exists between students and classes is represented by:

$$\frac{\tau_\pi}{\sigma^2 + \tau_\pi + \tau_\beta} = \frac{0.867}{1.698 + 0.867 + 0.984} = 0.244 \ (n=933)$$

At Level 3, or macro level, the proportion of the variance in the Chinese achievement that exists between classes is given by:

$$\frac{\tau_\beta}{\sigma^2 + \tau_\pi + \tau_\beta} = \frac{0.9840}{1.698 + 0.867 + 0.984} = 0.277 \ (n=39)$$

The above estimations have shown that the Level 1 or micro level model accounts for 48 per cent of the variance, the Level 2 or meso level model accounts for 24 per cent of the variance, and the Level 3 or macro model accounts for 28 per cent of the variance in the outcome measure of the Chinese language achievement under examination in this study. It is then possible to conclude that most of the variance exists within students (48%), followed by between-class variance (28%). A slightly smaller proportion of variance (24%) exists between students. Furthermore, the χ^2 statistics associated with these variance components recorded in Table 10.2 indicate highly significant variation between students, and between classes.

The three-level fully unconditional model also estimates the reliabilities at the between-student or meso level, and the between-class or macro level under investigation in this study. The reliabilities calculated are for the estimates of random Level 1 and Level 2 coefficients. Table 3 shows the reliabilities of the coefficients and indicates that the average reliability of the between-student sample mean for use in discrimination among students within the same class, P_0 is 0.616 at Level 2, and the reliability of the class sample mean as an estimate of the class mean, B_{00} is 0.928 at Level 3.

The final estimation of the fixed effects for the fully unconditional model is shown in Table 4. It should be noted that there is only one fixed effect in the fully unconditional model. In the present example, the fixed effect of G_{000} which indicates the grand mean test score in the Chinese language achievement was estimated to be 1.370 with a standard error of 0.165. This value is tested for a difference from zero with a *t*-value of 8.31. The above estimates obtained from the fully unconditional model are then used to determine the amount of variance explained by the final three-level model.

Table 3   Reliabilities of Random Level-1 and Level-2 Coefficients

Random Coefficient	Reliability Estimate
INTRCPT1, P_0 (Level 1 coefficient)	0.616
INTRCPT1/INTRCPT2, B_{00} (Level 2 coefficient)	0.928

Table 4   Final Estimation of Fixed Effects in the Fully Unconditional HLM Model

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
For INTRCPT1, P_0				
For INTRCPT2, B_{00}				
INTRCPT3, G_{000}	1.370	0.165	8.308	0.000

The final model

The final model for the three levels is written as Equations 4 to 15. The within-student equation in the three-level model is given as:

$$Y = P_0 + P_1 * (TIMEID) + E \qquad \text{[Equation 4]}$$

where

Y = the Rasch test score or outcome variable *score*;

P_0 = the intercept for the expected outcome variable *score*;

P_1 = the regression slope, namely the expected change in the Chinese language achievement associated with *TIMEID*; and

E = the random error at the micro or within-student level.

In words, Equation 4 states that the expected achievement in the Chinese language of a particular student depends upon the different points of time, namely Term 1, Term 2, Term 3 and Term 4.

For the Level 2 model, or meso level, there are seven hypothesised explanatory variables that are employed to investigate the variation in growth parameters between students. Equation 5 indicates that the predicted achievement of a student in learning the Chinese language depends on the gender, the amount of homework, the encouragement from parents, English word knowledge, effort in school work, perception of ease or difficulty of the Chinese language, and interest in learning the Chinese language. The Level 2 model is specified as:

$$
\begin{aligned}
P_0 = {} & B_{00} + B_{01} * (SEX) + B_{02} * (HOMEWRK1) + B_{03} * (ENCOURAG) \\
& + B_{04} * (SCOKWN) + B_{05} * (SCHEFFT1) + B_{06} * (CHIEASE1) \\
& + B_{07} * (INTREST2) + R_0 \qquad \text{[Equation 5]}
\end{aligned}
$$

$$P_1 = B_{10} + R_1 \qquad \text{[Equation 6]}$$

where

P_0 = the intercept, namely the expected outcome for each individual student;

B_{00} = the level-2 intercept ;

B_{01} = the regression slope associated with *SEX;*

B_{02} = the regression slope associated with *HOMEWRK1;*

B_{03} = the regression slope associated with *ENCOURAG* (i.e. the encouragement from parents);

B_{04} = the regression slope associated with *SCOKNW* (i.e. English word knowledge test scores);

B_{05} = the regression slope with *SCHEFFT1* (i.e. attitude towards and effort in school work);

B_{06} = the regression slope with *CHIEASE1* (i.e. ease or difficulty of the Chinese language);

B_{07} = the regression slope with *INTREST2* (i.e. interest in learning the Chinese language expressed in the second Attitude Questionnaire);

R_0 = the Level-2 random effect.

P_1 = the regression slope, namely the expected change in the Chinese language achievement associated with *TIMEID*;

B_{10} = the effect of occasion or time on the outcome variable *SCORE*;

R_1 = the regression slope associated with *TIMEID*.

At Level 3, namely between-class level in the final model, there are three hypothesised variables which are aggregated and averaged from Level 2 to examine the variation between classes. These three variables are *LEVEL_1*, *FOCC_1* and *EXPECT_1*. The Level 3 regression equations are described as follows:

$$B_{00} = G_{000} + G_{001} (LEVEL\_1) + U_{00} \qquad \text{[Equation 7]}$$

$$B_{01} = G_{010} + U_{01} \qquad \text{[Equation 8]}$$

$$B_{02} = G_{020} + G_{021} (FOCC\_1) + U_{02} \qquad \text{[Equation 9]}$$

$$B_{03} = G_{030} + U_{03} \qquad \text{[Equation 10]}$$

$$B_{04} = G_{040} + U_{04} \qquad \text{[Equation 11]}$$

$$B_{05} = G_{050} + U_{05} \qquad \text{[Equation 12]}$$

$$B_{06} = G_{060} + G_{061} (EXPECT\_1) + U_{06} \qquad \text{[Equation 13]}$$

$$B_{07} = G_{070} + U_{07} \qquad \text{[Equation 14]}$$

$$B_{10} = G_{100} + U_{10} \qquad \text{[Equation 15]}$$

where

B_{00} = the mean initial status;

G_{000} = the overall mean initial status;

G_{010} = the regression slope with *LEVEL_1* (the class level);

U_{00} = the class random effect associated with class;

$B_{01} \dots B_{07}$ = the regression slopes associated with *SEX, HOMEWRK1, ENCOURAG, SCOKNW, SCHEFFT1, CHIEASE1* and *INTREST2* respectively;

B_{10} = the effect of occasion or time on the outcome variable *SCORE*;

$G_{020} \dots G_{070}$ = the Level 3 intercepts;

G_{021} = the regression slope associated with *FOCC_1* (fathers' average occupation at the class level);

G_{061} = the regression slope associated with *EXPECT_1* (students' average expected occupation at the class level);

$U_{01} \dots U_{07}$ = the Level 3 random effects;

U_{10} = the class *SCORE* random effect;

G_{100} = the overall mean *SCORE* effect.

Table 5 presents the results obtained from the final model of the three-level HLM analysis for the estimates of the variance components. As a result of addition of predictors of the Chinese language achievement at all three levels, it can be seen that the estimates of variance components of three levels in the final model have decreased in comparison with those in the fully unconditional model. The proportion of variance explained by the final model at each level are computed by using the following three formulae:

within-students $\dfrac{\sigma^2_{(fully\ uncon.)} - \sigma^2_{(final)}}{\sigma^2_{(fully\ uncon.)}} = \dfrac{1.698 - 1.118}{1.698} = \dfrac{0.581}{1.698} = 0.34$

between-students $\dfrac{\tau_{\pi(fully\ uncon.)} - \tau_{\pi(final)}}{\tau_{\pi(fully\ uncon.)}} = \dfrac{0.867 - 0.661}{0.867} = \dfrac{0.207}{0.867} = 0.24$

between-classes $\dfrac{\tau_{\beta(fully\ uncon.)} - \tau_{\beta(final)}}{\tau_{\beta(fully\ uncon.)}} = \dfrac{0.984 - 0.195}{0.984} = \dfrac{0.789}{0.984} = 0.80$

Table 5   Final Estimation of Variance Components by Final Model

(a) Final estimation of Level 1 and Level 2 variance components

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P- value
INTRCPT1, R_0	0.813	0.661	588	1828.328	0.000
TIMEID slope, R_1	0.050	0.003	861	755.451	>.500
Level-1, E	1.057	1.118			

(b) Final estimation of level-3 variance components

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1/INTRCPT2,	U_{00}	0.442	0.195	29	71.263	0.000
INTRCPT1/ SEX,	U_{01}	0.118	0.014	30	39.425	0.116
INTRCPT1/HOMEWRK1,	U_{02}	0.200	0.040	29	33.764	0.248
INTRCPT1/ENCOURAG,	U_{03}	0.144	0.021	30	30.256	0.453
INTRCPT1/ SCOKNW,	U_{04}	0.126	0.016	30	31.956	0.369
INTRCPT1/SCHEFFT1,	U_{05}	0.050	0.003	30	22.801	>.500
INTRCPT1/CHIEASE1,	U_{06}	0.086	0.007	29	25.563	>.500
INTRCPT1/INTREST2,	U_{07}	0.160	0.026	30	32.254	0.356
TIMEID/INTRCPT2,	U_{10}	0.429	0.184	30	426.860	0.000

Notes: $E = \sigma^2$ $R_0 = \tau_\pi$ $U_{00} = \tau_\beta$

The model explains variance at all the three levels. The variance explained at Level 1 or the within-students level is 34 per cent. Only 24 per cent of the variance is explained at Level 2 or the between-students level. It is of interest to note that the variance explained at Level 3 or the between-classes level is 80 per cent. Once variance is calculated for each level, the calculation of total amount of variance explained by the model is possible. The total variance explained by the model is calculated by adding the total variance of each level obtained by multiplying variance explained by the final model and that predicted by the fully unconditional model. Thus, the proportion of total variance explained by the model is 44 per cent. The estimates employed in the calculation are given by:

within-students $= 0.34(final)$ x $0.48(fully\ uncon.) = 0.16$ (n=3,109)
between-students$= 0.24(final)$ x $0.24(fully\ uncon.) = 0.06$ (n=933)
between-classes $= 0.80(final)$ x $0.28(fully\ uncon.) = 0.22$ (n=39)

Total variance explained by the model $=0.16 + 0.06 + 0.22 = 0.44$

Table 6 records the reliabilities at the between-students and the between-classes levels for the final model. The figures indicate that the average reliability of the between-students intercept is approximately 0.60, and the reliability of the between-classes intercept is 0.52. The low level of reliability of *TIMEID* (0.008) should be noted, since it indicates marked variability in student performance overtime.

Final Estimation of Fixed Effects

The above sections address the estimates of variance components for the three levels and reliabilities for between-students and between-classes in the final model. The section that follows discusses the final estimation of fixed effects for the three levels in the final HLM model. Table 7 below presents the details.

Table 6   Reliabilities of Random Level 1 and Level 2 Coefficients by Final Model

Random Level 1 and Level 2 coefficient	Reliability estimate
INTRCPT1, P_0	0.595
TIMEID, P_1	0.008
INTRCPT1/INTRCPT2, B_{00}	0.519
INTRCPT1/ SEX, B_{01}	0.052
INTRCPT1/HOMEWRK1, B_{02}	0.284
INTRCPT1/ENCOURAG, B_{03}	0.136
INTRCPT1/ SCOKNW, B_{04}	0.144
INTRCPT1/SCHEFFT1, B_{05}	0.061
INTRCPT1/CHIEASE1, B_{06}	0.073
INTRCPT1/INTREST2, B_{07}	0.249
TIMEID/INTRCPT2, B_{10}	0.911

Level 3 Predictors

The first panel in Table 7 shows that the positive coefficient of the Level 3 predictor *level-1* is 0.17. This indicates that students in classes at higher year levels in learning the Chinese language are higher-level achievers in the Chinese language tests, with an increment of 0.17 logits for each year of learning after other factors have been taken into account.

Level 2 Predictors

At Level 2, namely the between-student level, a number of variables are observed to contribute to the Chinese language achievement scores across year levels and across occasions. First, sex of students (*sex*) shows a significant influence on achievement scores in the Chinese language tests with a coefficient of 0.25 logits. This indicates that girls are likely to be higher achievers than are boys on the Chinese language tests. The strongest predictor at this level is the score obtained on the English word knowledge tests (*scoknw*) with a coefficient of 0.46. This large positive coefficient indicates that students who score at a high level on the English word knowledge tests also achieve a higher score on the Chinese achievement tests. In other words, students' average English word knowledge proficiency has a direct effect on their achievement in the Chinese language: the better students are on the English word knowledge tests, the higher level they achieve on the Chinese language tests.

Table 7   Final Estimation of Fixed Effects by Final HLM Model (with robust standard errors)

Fixed Effect	Coefficient	StError	T-ratio	d.f.	P-value
For INTRCPT1, P_0					
For INTRCPT2, B_{00}					
INTRCPT3, G_{000}	1.222316	0.075527	16.184	37	0.000
LEVEL 1, G_{001}	0.166922	0.041641	4.009	37	0.000
For SEX, B_{01}					
INTRCPT 3, G_{010}	0.246229	0.069081	3.564	38	0.001
For HOMEWRK1, B_{02}					
INTRCPT3, G_{020}	0.110409	0.049273	2.241	37	0.031
FOCC_1, G_{021}	0.268258	0.084660	3.169	37	0.003
For ENCOURAG, B_{03}					
INTRCPT3, G_{030}	0.099839	0.050835	1.964	38	0.056
For SCOKNW, B_{04}					
INTRCPT3, G_{040}	0.464115	0.047205	9.832	38	0.000
For SCHEFFT1, B_{05}					
INTRCPT3, G_{050}	0.039159	0.019493	2.009	38	0.051
For CHIEASE1, B_{06}					
INTRCPT3, G_{060}	0.103101	0.051777	1.991	37	0.053
EXPECT_1, G_{061}	0.255907	0.083102	3.079	37	0.004
For INTREST2, B_{07}					
INTRCPT3, G_{070}	0.171162	0.042181	4.058	38	0.000
For TIMEID slope, P_1					
For INTRCPT2, B_{10}					
INTRCPT3, G_{100}	0.486559	0.072509	6.710	38	0.000

Note: Deviance = 10367.80 with 62 parameters

A third predictor is students' average interest in learning the Chinese language expressed in responding to the attitude questionnaire on the second occasion (*Intrest2*) with a positive coefficient of 0.17. This indicates that the more interested students are in learning the Chinese language at Term 4, the better they perform on the Chinese language tests. Thus, interest in learning the Chinese language is associated with the level of achievement on the tests.

Another variable that has an effect on *score* at Level 2 is encouragement from parents (*encourag*) with a coefficient of 0.10. This significant result shows that the more encouragement students receive from parents in learning the Chinese language, the higher score they are likely to achieve on the Chinese language tests. The positive coefficients of 0.11 for predictor *homewrk1* (homework in all school subjects) and 0.10 for predictor *chiease1* (ease or difficulty of learning the Chinese language at Term 1) indicate that those students who do more homework are likely to achieve higher scores on the Chines language tests, and those students who achieve high scores consider the Chinese language easier to learn.

The small but significant positive coefficient (y=0.04) associated with *schefft1* (effort in the schoolwork) indicates there is a relationship between students' effort in schoolwork as a whole

and their score in the Chinese language tests. The more effort a student puts into schoolwork, the higher score he or she is likely to achieve on the Chinese language tests.

Level 1 Predictors

The Level 1 predictor *timeid* is found to have an impact on *score.* The large coefficient of 0.49 indicates that the performance of students improves markedly across the terms of a school year. Students gain approximately 0.5 logits in the interval from Term 1 to Terms 2, and from Term 2 to Term 3, and from Term 3 to Term 4. There is clearly, in general, very effective learning occurring during the four terms of the 1999 school year, particularly at the lower grades where there are large numbers of students.

Furthermore, the sizeable magnitude of the Level 1 predictor coefficient (0.49) of nearly half a logit per term on average compared with the low magnitude of the Level 3 effect of variable *level_1* (0.17) warrants careful consideration. This growth in learning the Chinese language occurs mainly between terms while the gain between levels may merely reflect the dropout of lower performing students between grade levels. It should be noted that these effects are estimated after the substantial effects of growth in English word knowledge are taken into account. There is no interaction effect observed between Level 2 variables and Level 1 variable *timeid*, namely between the between-student and the within-student levels.

In addition, from the null model (Deviance = 11489.25 with 4 parameters) there is a significant reduction in deviance for the final model (Deviance = 10367.80 with 61 parameters), indicating that the final model is a much better model than the null model in fitting the data.

The final three-level HLM model for the Chinese language achievement is presented diagrammatically in Figure 1 using the information in Table 7, where the robust regression effects are recorded. The interaction effects are best shown by calculating the co-ordinates for a graphical representation, and the procedures for the calculations are described below.

It should be noted that in the presentation of results, the robust regression solution has been presented because there is clearly heterogeneity of variance in the criterion measure across grades and the possible presence of some outliers in the data. Consequently, the robust regression results would seem to provide a more stable and meaningful analysis of the data collected.

Interaction Effects

In addition to the predictors that show direct influence on the scores of the Chinese language tests, Table 7 also reports two interaction effects between Level 3 and Level 2 predictors of the outcome variable *score*. The two interaction effects involve class average occupation of fathers (*focc-1*) on homework in all school subjects (*homewrk1*), and class average expected occupation of students (*expect-1*) on students' perception of ease or difficulty of the Chinese language (*chiease1*). The coefficient of *focc-1* (Level 3 variable) on *homewrk1* (Level 2 predictor) (y=0.27, t=2.6) indicates a positive relationship between class average of father's occupation and the effect that amount of homework has on the outcome variable *score.* This result indicates that students in those classes where fathers have a higher level of occupation and who do more homework than students in classes whose fathers have a lower level occupation achieve higher scores on the Chinese language achievement tests.

The other interaction effect is found between *expect_1* (Level 3 variable) and *chiease1* (Level 2 predictor) (y=0.26). This result indicates that the interactive effect of class average expected occupation between high- and low-viewers of the ease of learning the Chinese language has an impact on scores on the Chinese language tests. Those students who are in classes that have a

higher level of expected occupation in the future and who find it easy to learn the Chinese language achieve higher scores on the Chinese language tests than those students in classes where there is a lower level of expected occupation. The nature of these interaction effects is examined more thoroughly in later sections.

Calculation of Co-ordinates for Interaction Effects

The calculations are based on the results of the final estimation of the fixed effects from the three-level HLM analysis presented in Table 6. The calculation for the co-ordinates of the graphs follows the procedures that Lietz (1995) employed for a two-level HLM analysis of reading achievement. Mohandas (1999) used the same procedures for a two-level HLM analysis of mathematics and science achievement.

Interaction Effects between Father's Occupation (focc-1) and Homework (homewrk1)

First, consideration is given to the interaction effect of the Level 3 variable *focc-1* with the Level 2 predictor *homewrk1* on achievement in learning the Chinese language. Figure 1 shows that *focc-1* as a Level 3 predictor has an interaction effect on the path between *homewrk1* as a Level 2 predictor and the outcome variable *score*. The equation for the Level 2 (i.e. between-student) predictor *homewrk1* to influence the outcome variable *score* is given below based on Equation 5 for the Level 2 model in the final model specification.

$$P_0 = B_{00} + B_{02}*(HOMEWRK1) + R_0 \qquad\qquad \text{[Equation 16]}$$

where

P_0 = the intercept for the expected outcome variable *SCORE*;

B_{00} = the Level 2 intercept;

B_{02} = regression slope associated with *HOMEWRK1* (i.e. homework for all school subjects);

R_0 = the Level 2 random effect.

The equation for the Level 3 (between classes) variable *focc-1* to have an interaction effect with the variable *homewrk1* on *score* is given in Equation 17 that is based on Equation 9 in the final model specification. It should be noted that at this level the variable *focc-1* involves the average occupational status of the fathers of students within a class.

$$B_{02}\ \ = G_{020} + G_{021}\ (FOCC\text{-}1) + U_{02} \qquad\qquad \text{[Equation 17]}$$

where

B_{02} = the regression slope associated with *HOMEWRK1*;

G_{020} = the Level 3 intercept;

G_{021} = the regression slope with *FOCC-1* (i.e. father's occupation);

U_{02} = the Level 3 random effect.

When B_{02} in Equation 17 is substituted in Equation 16, Equation 18 is obtained:

$$P_0 = B_{00} + G_{020}*(HOMEWRK1) + G_{021}*(FOCC\text{-}1)*(HOMEWRK1) + R_0 \quad \text{[Equation 18]}$$
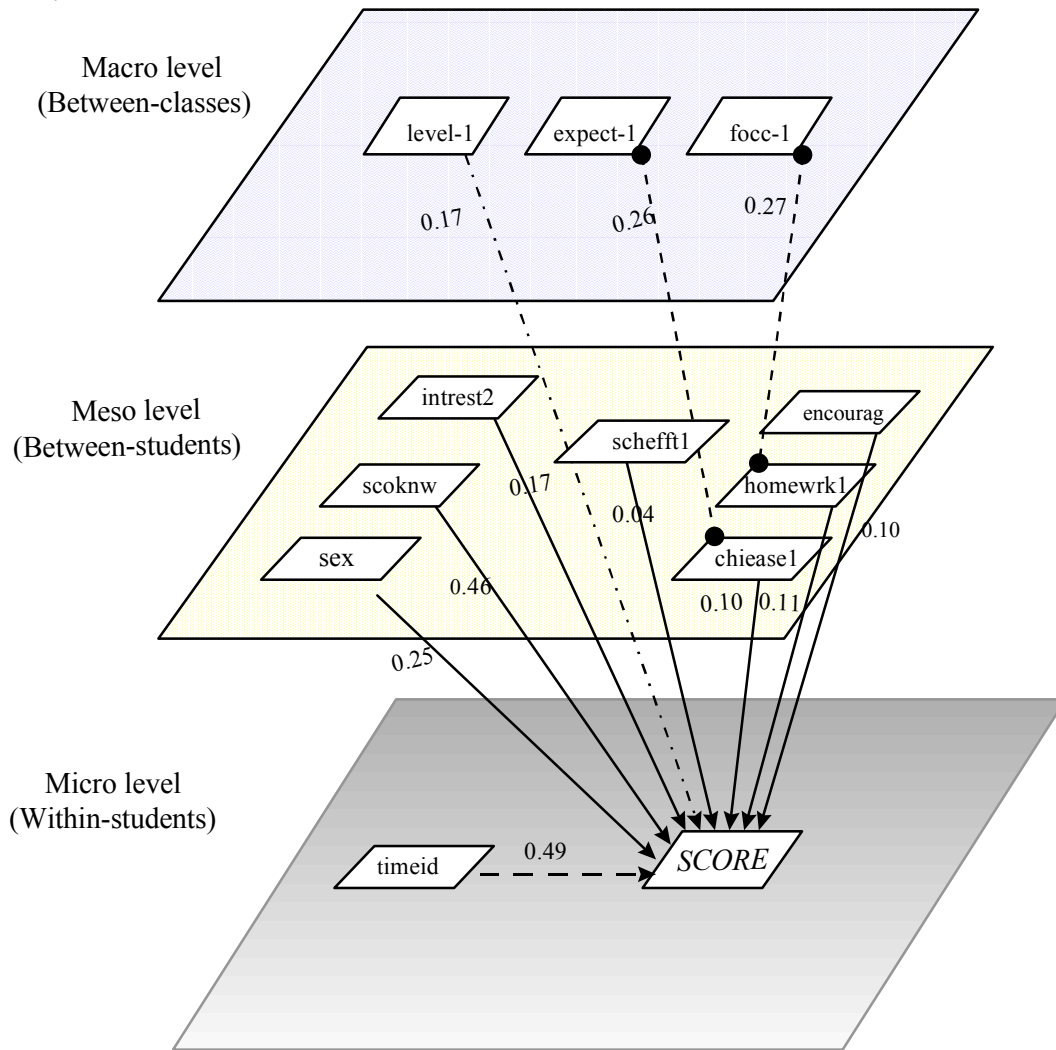
where (see Table 6)

$B_{00} = G_{000}$ (as the overall mean of Chinese achievement across classes) = 1.22;

$G_{020} = 0.11$; and

$G_{021} = 0.27$

**Figure 2   Final Three-level Hierarchical Linear Model of Factors Influencing
Achievement in Learning the Chinese Language**

When the coefficients of G_{000}, G_{020}, and G_{021}, namely, 1.22, 0.11, and 0.27 respectively, are
substituted in Equation 18, Equation 19 is obtained:

$P_0 = 1.22 + 0.11*(HOMEWRK1) + 0.27* (FOCC\text{-}1) *(HOMEWRK1) + R_0$ [Equation 19]

Co-ordinates are calculated for points, one score point above and below the variable means and at the mean value.

(1). Higher father's occupation and more homework (*FOCC-1*=1; *HOMEWRK1*=1)
P_0 (*SCORE*) = 1.22 + 0.11* (1) + 0.27 * (1)*(1) = 1.60

(2). Higher father's occupation and less homework (*FOCC-1*=1; *HOMEWRK1*=-1)
P_0 (*SCORE*) = 1.22 + 0.11*(-1) + 0.27 *(1)*(-1) = 0.84

(3). Lower father's occupation and more homework (*FOCC-1*= -1; *HOMEWRK1=1*)
P_0 (*SCORE*) = 1.22 + 0.11*(1) + 0.27*(-1)*(1) = 1.06

(4). Lower father's occupation and less homework (*FOCC-1*=-1; *HOMEWRK1*=-1)
P_0 (*SCORE*) = 1.22 + 0.11*(-1) +0.27*(-1)*(-1) = 1.38

(5). Average father's occupation and more homework (*FOCC-1*=0; *HOMEWRK1*=1)
P_0 (*SCORE*) = 1.22 + 0.11*(1) + 0.27*(0)*(1) = 1.33

(6). Average father's occupation and less homework (*FOCC-1*=0; *HOMEWRK1*= -1)
P_0 (*SCORE*) = 1.22 + 0.11*(-1) +0.27*(0)*(-1) = 1.11

The coordinates obtained from the above calculations are used to produce Figure 3 which shows the interaction effect of class average father's occupation with the amount of school homework on achievement in learning the Chinese language. The regression lines indicate that :

(a). those students in classes whose fathers on average have higher level occupations and who do more homework, achieve higher scores on the Chinese language tests which is shown in Figure 2,
(b). those students in classes whose fathers on average have higher level occupations and who do less homework achieve lower scores on the Chinese language tests,
(c). those students in classes whose fathers on average have lower level occupations and who do less homework are relatively high achievers on the Chinese language tests, and
(d). those students in classes whose fathers on average have lower occupations and who do more homework are relatively low achievers on the Chinese language tests.

The result indicates that those students whose fathers have a higher level occupation are likely to gain a higher level of achievement in learning the Chinese language than do the students whose fathers work in a lower level of occupation (see Figure 3).

Interaction Effects between Students' Expected Occupation (expect-1) and Ease of Learning the Chinese Language (chiease1)

The same procedures can be employed to show the interaction effect between the average level within a class of a student's expected occupation (*expect-1*) and the student's perception about ease of learning the Chinese language (*chiease1*) on achievement in learning the Chinese language (*score*) based on Equation 5 for the Level 2 model in the final model specification.

$$P_0 = B_{00} + B_{06}*(\textbf{\textit{CHIEASE1}}) + R_0 \qquad \text{[Equation 20]}$$
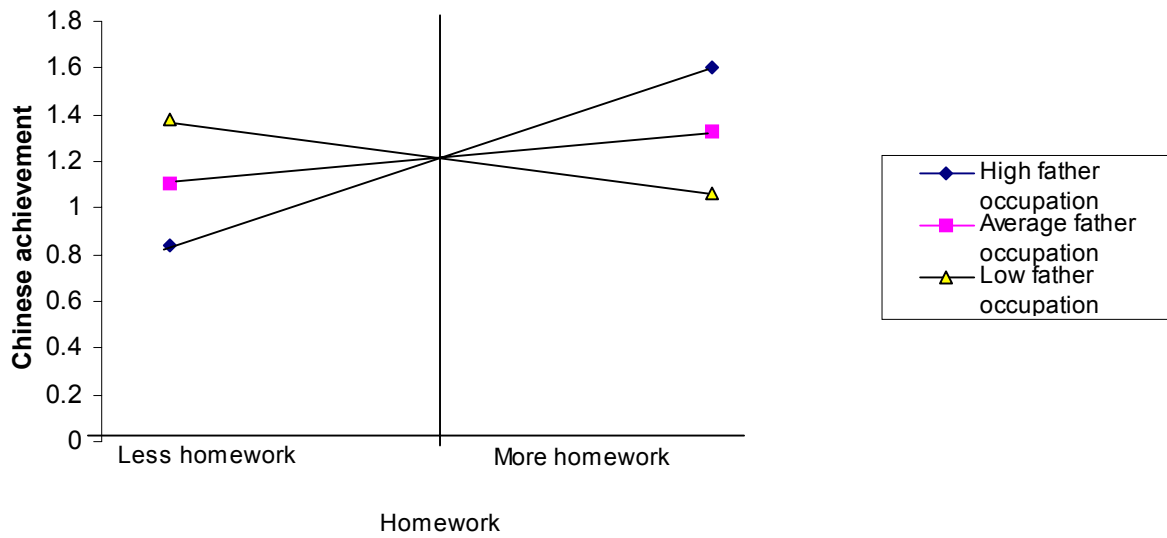
where

P_0 = the intercept for the expected outcome variable *SCORE*;

B_{00} = the Level 2 intercept;

B_{06} = the regression slope associated with *CHIEASE1* (i.e. student's perception about ease of learning the Chinese language);

R_0 = the Level 2 random effect.

Figure 3   **Impact of the Interaction Effect of Father's Occupation with the Amount of Homework on Chinese Achievement**

The equation for the Level 3 (between classes) variable *expect-1* to have an interaction effect with the Level 2 variable *chiease1* on *score* is given below based on Equation 13 in the final model specification.

$$B_{06} = G_{060} + G_{061} (EXPECT\text{-}1) + U_{06}$$
[Equation 21]

where

B_{06} = regression slope associated with *CHIEASE1*;

G_{060} = level 3 intercept;

G_{061} = regression slope with *EXPECT-1* (i.e. average level of student's expectations in class);

U_{06} = Level 3 random effect.

When B_{06} in Equation 21 is substituted in Equation 20, Equation 22 is obtained:

$$P_0 = B_{00} + G_{060}*(CHIEASE1) + G_{061}*(EXPECT\text{-}1)*(CHIEASE1) + R_0$$
[Equation 22]

where (see Table 6)

$B_{00} = G_{000} = 1.22$;

$G_{060} = 0.10$; and

$G_{061} = 0.26$

When these values (1.22, 0.10, and 0.26) are substituted in Equation 22:

$$P_0 = 1.22 + 0.10*(CHIEASE1) + 0.26*(EXPECT\text{-}1)*(CHIEASE1) + R_0$$
[Equation 23]

The coordinates for the graphical representation are calculated for points one score point above and below the variable means and at the mean value.

(1). High expected occupations and high ease of learning the Chinese language (*EXPECT-1*=1; *CHIEASE1*=1)

P_0 (*SCORE*) = 1.22 + 0.10*(1) +0.26*(1)*(1) = 1.58

(2). High expected occupations and low ease of learning the Chinese language (*EXPECT-1*=1; *CHIEASE1*=-1)

P_0 (*SCORE*) = 1.22 + 0.10*(-1) +0.26*(1)*(-1) = 0.86

(3). Low expected occupations and high ease of learning the Chinese language (*EXPECT-1*= -1; *CHIEASE1*=1)

P_0 (*SCORE*) = 1.22 + 0.10*(1) +0.26*(-1)*(1) = 1.06

(4). Low expected occupations and low ease of learning the Chinese language (*EXPECT-1*= -1; *CHIEASE1*= -1)

P_0 (*SCORE)* = 1.22 + 0.10*(-1) +0.26*(-1)*(-1) = 1.38

(5). Average expected occupation and high ease of learning the Chinese language (*EXPECT-1*=0; *CHIEASE1*=1)

P_0 (*SCORE*) = 1.22 + 0.10*(1) +0.26*(0)*(1) = 1.32

(6). Average expected occupation and low ease of learning the Chinese language (*EXPECT-1*=0; *CHIEASE1*=-1)

P_0 (*SCORE*) = 1.22 + 0.10*(-1) +0.26*(0)*(-1) = 1.12

The coordinates obtained from the above calculations are used to produce Figure 4 which shows the interaction effect of the average level of a student's expected occupations with the student's perception about the ease or difficulty of learning the Chinese language on the student's Chinese language test scores. The regression lines indicate, as shown in Figure 4, that:

(a). those students in classes which who have on average a higher level of expected occupation and who find that it is easy to learn the Chinese language achieve higher scores on the Chinese language tests,

(b). those students in classes which have on average a lower level of expected occupation and who have more difficulty in learning the Chinese language achieve relatively high scores on the Chinese language tests,

(c). those students in classes which have on average a higher level of expected occupation and who have more difficulty in learning the Chinese language achieve lower scores on the Chinese language tests, and

(d). those students in classes which have on average a lower level of expected occupation and who find the Chinese language easy to learn achieve relatively low scores on the Chinese language tests.

CONCLUSION

The findings of the three-level HLM analysis of factors influencing achievement *score* on the Chinese language tests are discussed in the above sections. At the within-student level*, timeid* is found to have an influence on the *score*. Seven variables at Level 2, namely the between-student level, have significant influences on *score* on the Chinese language tests. These seven variables are sex of students (*sex*), score for English Word Knowledge (*scoknw*), interest expressed in learning the Chinese language on the second occasion *(intrest2)*, encouragement from parents (*encourag*), homework for all school subjects (*homewrk1*), effort given to school work indicated on Occasion 1 (s*cheff1*), and perception about ease or difficulty of learning the Chinese language indicated on Occasion 1(*chiease1*). At Level 3, namely the between-class level, grade level (*Level

1) is observed to have an influence on achievement score. The positive effect indicates that the higher the grade level, the greater the scores the students achieved on the Chinese language tests.
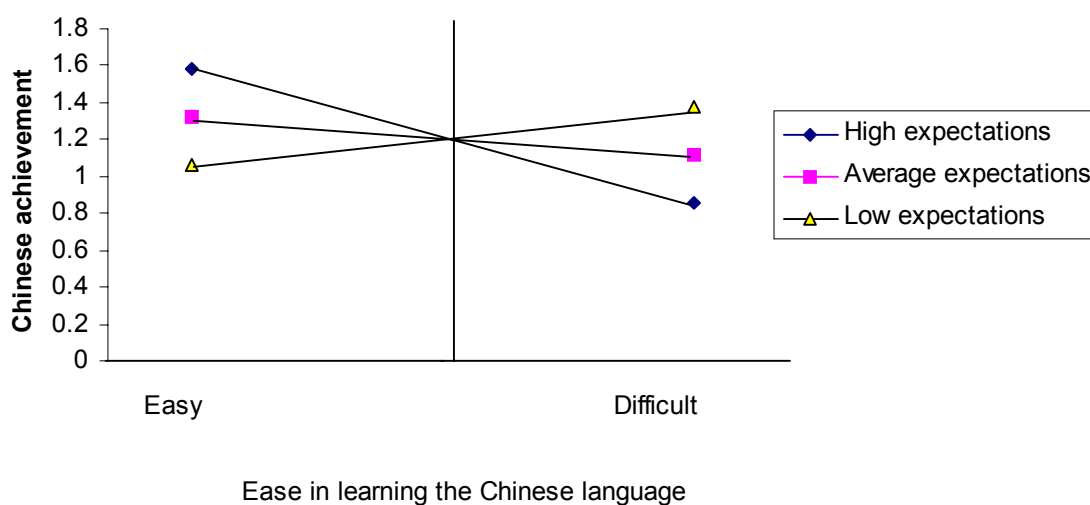
Figure 4   Impact of the Interaction Effect of Students' Expected Occupation on Ease of Learning the Chinese Language on Chinese Achievement

Apart from these effects, two Level 3 variables, that is, class average of students' expected occupations (*expect-1*) and class average of father's occupation (*focc-1*) show interaction effects with Level 2 predictors *chiease1* and *homewrk1* respectively on the outcome variable *score* on the Chinese language tests.

In summary, an examination has been made of factors that influence student achievement in learning the Chinese language as a school subject by employing a three-level HLM analysis. These findings, particularly those involving homework and attitudes have important implications for the learning of the Chinese language in Australian schools.

REFERENCES

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. The Flinders University of South Australia, School of Education.

Andrich, D. (1988). *Rasch models for measurement.* Series: Quantitative applications in the social sciences. Sage Publications, Inc.

Bryk, A. S. & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level linear model. *American Journal of Education*, 97 (1): 65-108.

Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage Publications.

Burstall, C. M. (1975). Factors affecting foreign-language learning: a consideration of some recent research findings. *Language Learning, 8, 1-21.*

Carroll, J. B. (1963a). A model of school learning. *Teachers College Record*, 64, 723-733.

Carroll, J. b. (1967). *The foreign language attainments of language majors in the senior year: a survey conducted in U.S. colleges and universities.* Cambridge, Mass: Laboratory for Research in Instruction, Graduate School of Education, Harvard University.

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.

Cheung, K. C.; Keeves, J. P.; Sellin, N. and Tsoi, S. C. (1990). The analysis of multilevel data in educational research: Studies of problems and their solutions. *International Journal of Educational Research.* Vol. 14:3, 215-319.

Goldstein, H. (1987). *Multilevel Models in educational and social research.* Oxford University Press, Oxford.

Hox, J. J. (1994). *Applied multilevel analysis.* TT-Publikaties, Amsterdam.

Hyltenstam, K. and Pienemann, M. (Eds.) (1985). *Modeling and assessing second language acquisition.* Multilingual Matters Ltd.

Johnson, T. G. (1995). *Multilevel models of student achievement in science.* Unpublished doctoral thesis, The Flinders University of South Australia, Adelaide, Australia.

Keeves, J. P. (1972). *Educational environment and student achievement:* A Multivariate study of the contributions of the home, the school and the peer group to change in mathematics and science performance during the first year at secondary school. ACER, Hawthorn, Victoria.

Keeves, J. P. and Sellin, N. (1992). Multilevel analysis. In: J. P. Keeves (ed.), *The IEA Technical Handbook*, The Hague, The Netherlands: IEA, pp. 307-326

Keeves, J. P. (1997). Models and model building. In J. P. Keeves (ed.): *Educational Research, Methodology, and Measurement: An International Handbook,* 2nd edn.). Cambridge, UK: Pergamon Press, pp.386-394.

Lietz, P. (1995). *Changes in reading comprehension across cultures and over time.* Unpublished Doctoral Thesis. The Flinders University of South Australia, School of Education.

Mason, W. M.; Wong, G. Y. and Entwistle, B. (1983). Contextual analysis through the multilevel linear model. In: Leinhardt, S. (ed.) (1983), *Sociological Methodology 1983-1984,* Vol. 14. Jossey-Bass, San Francisco, California.

McNamara, T. F. (1996). *Measuring second language performance.* London: Longman.

Mohandas, R. (1999). *Mathematics and science achievement of junior secondary school students in Indonesia.* Unpublished Doctoral Thesis. The Flinders University of South Australia, School of Education.

Prosser, R.; Rasbash, J. and Goldstein, H. (1991). *ML3 software for three-level analysis. Users Guide.* London: Institute of Education, University of London.

Raudenbush, S. W. and Bryk, A. S. (1994). Hierarchical Linear Model. In T. Husén and T.N. Postlethwaite (eds), *The International Encyclopedia of Education (2nd. Edn), pp. 2590-2596,* Oxford: Pergamon Press.

Raudenbush, S.; Bryk, T. and Congdon, R. (2000). *HLM 5, Scientific software international, Inc.*

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, Vol.32:3, pp. 583-625.