International Education Journal

Volume 2 Number 2
July 2001

Published by

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

International Education Journal

Volume 2, Number 2, July 2001

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

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

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

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.

The views and styles expressed in the articles in this publication are those of the individual authors and are not necessarily shared by the reviewers, the editors or the editorial advisory board.

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

Copyright © 2001

Shannon Research Press

ISSN 1443-1475

All rights reserved. Apart from individual use, no part of this publication may be reproduced in any form or by any means without prior written permission from the publisher.

Designed by Katherine Dix

Printed in Adelaide
March 2001

CONTENTS

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

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

Editorial Board

Editors in Chief

Dr Reaburn Reynolds
Flinders University (Education)
Bedford Park, South Australia
Phone: 08-8201 2807,
Fax: 08-8201 3184,
reaburn.reynolds@flinders.edu.au

Prof John Keeves
Flinders University (Education)
Bedford Park, South Australia
Phone: 08-8201 2392,
Fax: 08-8201 3184,
john.keeves@flinders.edu.au

Online Editor

Ms Katherine L. Dix
Flinders University (Education)
Bedford Park, South Australia
Phone: 08-8201 2105,
Fax: 08-8201 3184,
katherine.dix@flinders.edu.au

Editorial Advisory Board

Prof Jonathan Anderson
School of Education
Flinders University

Prof Bob Catley
Department of Political Studies
University of Otago

Prof Kevin Marjoribanks
Graduate School of Education
University of Adelaide

Dr Larry Saha
Department of Sociology
Australian National University

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.

Perspectives on Teacher Burnout and School Reform

A. Gary Dworkin
The University of Houston and The Australian National University *gdworkin@mail.uh.edu*

This paper examines two general models of teacher burnout, the psychological and the sociological. In the sociological model burnout is a form of job-specific alienation and can be redressed through organizational and structural changes. School reform in the United States as it impacts teachers has followed three waves since 1983: legislated standardization and competency testing, decentralization and site-based decision making, and high-stakes testing with accountability. Teacher burnout data collected during each of the waves is compared with data collected prior to the reforms. Each wave exacerbated teacher burnout, but affected different sub-groups of teachers.

teacher burnout, school reform, United States educational policies

INTRODUCTION

Imagine that there appeared an announcement in the want-ads section of the local newspaper: What would it say? How about this?

> Wanted, college-educated individuals who are willing to put in excessively long hours without commensurate compensation; who can work under adverse conditions, with unappreciative supervisors and even more unappreciative clients, many of whom prefer to be uninvolved, as well; who do not mind having inadequate resources and support services; who agree to assume unspecified responsibilities without prior notification; but who will be held accountable for the satisfaction and performance of the unappreciative and uninvolved clients. Candidates for the positions also must be willing to receive inadequate wages and expect not to be able to double their income in constant dollars in a lifetime. Applicants are encouraged to send resumes to the Teacher Employment Office of the _____School. (Adapted from a want ad idea of Linda Darling-Hammond 1983).

The hypothetical advertisement above depicts the working conditions of public school teachers, especially in urban school districts in the United States. Menlo and Poppleton (1990) suggest that similar conditions exist in most developed nations. It is only in the developing nations that public school teachers garner the respect they desire, in part because they represent part of the educated elite in those nations.

BURNOUT AS A PSYCHOLOGICAL AND SOCIOLOGICAL CONSTRUCT

Burnout as a construct emerged from the writings of the clinical psychologist, H. J. Freudenberger, beginning with his publication in the *Journal of Social Issues* in 1974. For Freudenberger, burnout represented a malaise of human service professionals, such as social workers, mental health workers, nurses, and teachers, that is characterized by feelings of "wearing out". Faced with a plethora of stressors on the job many human service professionals become emotionally exhausted and lose their sense of purpose or of accomplishment. By the late 1970s other psychologists

operationalized burnout in terms of three central dimensions: emotional exhaustion, loss of a sense of personal accomplishment, and depersonalization (Maslach 1978a, 1978b, 1993; Maslach and Jackson 1982, Cherniss 1980a, 1980b). In the Maslach et al. viewpoint burnout occurs when over-stressed individuals feel emotionally drained by their work environment, feel that their activities result in no benefit to those they had intended to help or to themselves, and come to blame their clients, patients, or students for failing to improve, get better, or learn, and in turn, for the professional's loss of feelings of accomplishment.

Other investigators, including Pines, Aronson, and Kafry (1981) described burnout as similar to tedium, but then retracted that view as implying that burnout was trivial (Pines and Aronson 1988). Later, Pines (1993) saw it as an existential crisis, linked to a sense of meaninglessness. That is, to the extent that professionals come to incorporate their work into their self-image, conditions that diminish the personal assessment of the value of that work likewise diminish the assessment of self-worth. When this happens, Pines argues, human service professionals come to ask, "Why am I doing what I am doing?" – a question reflecting self-doubt and a crisis of existence.

There is one constant element in the psychological perspectives. Because burnout is viewed as the inability of the individual to cope with stressors, the treatment of burnout lies in helping the individual learn to cope. Psychologists have urged the burned out to engage in a variety of clinical solutions from stress management (Cedoline 1982; Swick and Hanley 1983, Gold and Roth 1993, and Pines 1993) to holistic health practices (Tubesing and Tubesing 1982). Such practices as yoga, deep breathing exercises, naps, and mental health breaks are recommended by the psychologists who blame the victims of burnout for their maladaptive behaviours.

Sociologists see burnout differently. Burnout is conceptualized as a form of alienation involving the dimensions described by Seeman (1959, 1975), including powerlessness, meaninglessness, normlessness, isolation, and estrangement. Alienation is seen as having organizational and social structural roots and therefore should not be addressed by the teaching of coping skills, but rather through structural change. Stress can still be a precipitating factor in burnout, as it is in the psychological models, but the causal elements of burnout are to be seen within the structure of the school or the structure of the educational system (Dworkin 1987; LeCompte and Dworkin 1991; Dworkin and Townsend 1994). When professionals are unable to negotiate agreements on role performances or to determine what are the role expectations within a human service organization, they acquire a sense of powerlessness (Shinn 1982), which soon leads to a sense of meaninglessness. Soon too, the individual begins to withdraw from social relationships within the organization (isolation) and to question whether continued participation in the organizational role is consistent with their self-conception (estrangement). In addition, the burned out individuals begin to blame their clients, students, or patients for failing to improve. Some may even feel that their clients or students refuse to improve or learn specifically to "spite" the burned out professional. In addition, the burned out professionals often feel that the organization is characterized by a degree of normlessness. That is, they feel that either there are no rules or that following the rules tends to be dysfunctional. Sparks and Hammond (1981) reported that burned out professionals often report that the rules of the organization are either unenforceable or cannot be interpreted. Dworkin (1987:28), after reviewing the literature from both the psychological and sociological traditions incorporated all five elements of alienation into his definition of burnout.

A story about the differences between the psychological and sociological perspectives can be instructive. The author was invited to meet with a group of teachers whose school district had enrolled them in a health maintenance organization (HMO) in the Houston area. The topic was on coping with job stress and the HMO also invited a clinical psychologist to address the teachers. One young teacher in her first year in the classroom stood up to ask a question. She reported that

she had been having trouble with classroom management and had asked her principal for help. One day, the principal came to her room and led her class in jeering at her and calling her a "poor teacher". She reported that she broke into tears and fled the room. When asked to reply to the scenario the clinical psychologist suggested that she should improve her coping skills with yoga. As a sociologist who looks for systemic problems I recommended that there needed to be a change in principals, or at least a retraining of that principal. There is an economy of scale in the sociological suggestion. In a large school district with 12,000 teachers and 350 principals it is easier and more efficient to change the administrative behaviour of the principals than to have to provide coping skills each year to all of the teachers.

SOCIETAL AND SOCIAL STRUCTURAL FACTORS IN TEACHER BURNOUT

There are numerous structural and organizational factors involved in teacher burnout. To mention only a few one might cite the declining public confidence in public education that has tended to devalue the teaching professional, especially in advanced industrialized nations. Opinion polls over the past 30 to 40 years have shown that the public believes that schools are not performing anywhere as well as they did in the past. Often the competence of public school teachers is mentioned as a contributing factor.

As public confidence declined the willingness of the best and the brightest to elect careers in teaching also declined. National studies in the United States reveal that individuals choosing a major in education represent the lowest of the entrance test scorers among university students, and those remaining in careers in education are among the lower scorers within that group. Furthermore, as public school teaching is a career dominated by women the pool of potential teachers becomes restricted in terms of abilities as higher paying, more prestigious occupations become more available to women.

There was a time when teachers represented an intellectual elite in the United States and other economically advanced nations. Having finished high school and usually having completed at least two years of tertiary education in the so-called normal schools (teacher's schools), public school teachers in the early years of the twentieth century were far better educated than the parents of their students. In fact, fewer than ten per cent of the American adult population finished high school as of the first decades of the century and it was not until the early 1940s that half of the adult population had a high school diploma. Today the figure approaches 90 per cent and about one-third of all adults aged 25-29 in the United States have at least a college degree. In many schools, then, the parents are at least as well educated as the teachers, and often have better education. However, the parents often have degrees in majors known to require higher entrance scores than did the teachers. Many middle class parents today meet teachers in the parent-teacher conferences assuming that they are interacting with a less able individual than themselves and even offer to "teach" the teachers how to introduce subject matter. By contrast, in the 1940s the teacher was the skilled expert.

Yet another factor affecting teacher morale is the gap between the expectations created in pre-service training and the experiences of teachers in classrooms, especially the highly stressful classrooms of high-poverty schools. Pre-service public school teachers come to expect through their training that they will be accorded professional autonomy and professional respect. They often feel that teaching is a calling and that their students will eagerly accept the knowledge that they have to offer. Their experiences are at considerable odds with their expectations. They are often faced with few resources in the classroom and treated with little respect and much abuse. Studies by the National Center for Education Statistics recount abuse and attacks on teachers in many schools. Teachers are victims of thefts and physical attack. LeCompte and Dworkin (1991)

chronicle many of these abuses, as does the annual report from the United States Department of Education (*The Condition of Education*).

Finally, I turn to one underemphasized structural factor that affects teacher morale. This is the role of legislated school reform, especially those reforms that have followed the issuance of the report of a presidential commission in 1983. What follows in this paper is an examination of the effects of school reform on teacher burnout.

THREE WAVES OF PUBLIC SCHOOL REFORM IN THE UNITED STATES

School reform in the United States, as it affects public school teacher attitudes, has undergone three discernable waves since the publication of *A Nation at Risk* by the National Commission on Excellence in Education (1983). The report of the commission decried the poor academic performance of American school children and predicted that unless a major revamp of American public education did not occur quickly, America would soon no longer hold on to its economic position in the global economy. By the next year the United States Secretary of Education declared that a nation had responded to the challenge with sweeping reforms. The first wave of such reforms sought "…to introduce uniformity and conformity through standardized curricula, rigorous requirements for student performance, promotion and graduation, and teacher evaluation (Smylie and Denny 1990: 235). The goal of the first wave of reforms was to guarantee that only competent teachers were in the classroom and that only educated students graduated from school. The first wave of reforms could be the "legislated wave." Soon, however, it was discovered that the first wave of reforms had failed and that it was time for a new tack to be taken. A second wave of reforms followed in the late 1980s. If centralized, legislated reform did not work, then it was time for decentralization. Localized or site-based decision making characterized the second wave, but with decentralization came localized accountability. That is, if teachers and principals were to be given more authority and more autonomy from the central district, they should also be held accountable for student learning outcomes.

It was really not unexpected that these reforms would also fail to boost student achievement to the levels expected by government, the public, and corporate America. Thus, by the early to mid-1990s came the third wave of reforms. Sometimes referred to as "high-stakes testing," the third wave depended upon the use of state-mandated standardized achievement tests, systems of rating schools and school districts, and holding students, teachers, and school administrators accountable for the results of those tests. While achievement tests have been used in the United States for decades, dating back at least to the 1920s, the purpose of the tests throughout most of the century was to permit teachers to assess the deficiencies of individual children and to develop curricula to redress those deficiencies. In the case of high-stakes tests the results are used to offer or withhold diplomas and graduation from students, to assess the continued accreditation of schools and school districts, and determine which teachers should be permitted to retain their jobs and which should be terminated from employment.

First-wave reforms in Texas were legislated under House Bill 72, passed in 1984. The legislation mandated competency testing of teachers, regardless of their prior success in the classroom, and the termination of those teachers who could not pass the test, set a seventh-grade reading level, after repeated attempts. The legislation also established a career ladder with salary increments associated with each rung (however, the legislature never provided the funding to support the legislated pay levels), but did implement a system of in-class assessments for placement on the career ladder. The legislation further standardized grading in all schools (passing grades were set at 70 per cent), set maximum primary grade class sizes, and established the principle of "no pass, no

play," whereby students who failed courses were barred from participation in extracurricular activities. This was "Top-Down Reform".

Texas' second-wave reforms occurred in 1989. Weeding out of incompetent teachers had not produced substantial gains in student achievement or reductions in the dropout rate, especially among minority students and students from low-income families. If centralization was failing, then decentralization must be the answer. The legislature mandated that schools become autonomous, or else. Legislation ordered the creation of site-based decision-making committees at every school in the state. The committees usually consisted of the principal, key teachers, representatives of parents' groups, and community stakeholders. Frequently, however, there were "turf battles" among the participants over control of the school.

The third wave of reforms began with the advent of the Texas Assessment of Academic Skills (TAAS) test, a criterion-referenced test. The test has mathematics, reading, and writing sections. Implementation of the test began in October 1990 with administration of only a limited number of grade levels; further, successive grade levels were not regularly tested, thereby making the determination of academic progress over successive grades impossible. However, beginning in 1994, students in Grades 3 through 8 and Grade 10 were tested each year, thus permitting the determination of annual gain scores. The three sections of the test are administered only in Grades 4, 8, and 10, while the other grades take only the reading and mathematics tests. Tenth-grade students must pass reading, mathematics, and writing sections of the TAAS, as well as their course-exit exams, before graduating from high school. The TAAS is administered in Spanish for students enrolled in bilingual education programs in Grades 3 through 6, but students in the English as a second language program must take the TAAS in English.

TAAS performance is an integral component of the state's public school accountability system, with test passage affecting high school graduation of students and passage rates determining the rating of schools and districts. However, the test is not the only element in the accountability system, as dropout and attendance rates are incorporated into the overall ratings. Campus and district results are readily available to the public at both the Texas Education Agency's website and physical copies of ratings are required to be prominently displayed at school district offices. School districts and individual campuses with high TAAS passage rates are offered financial rewards, whereas low-performing campuses and districts may face take-over by the state, removal of administrators, and even loss of accreditation. The accountability results are disaggregated to the passage rates by minority and economically disadvantaged populations; this has the beneficial effect of causing districts not to disregard historically ignored categories of students.

If the first wave saw teachers as the problem, the second saw teachers as the solution, and the third saw all participants in schooling as problematic. Clearly each of the waves had effects on the morale of teachers; each wave altered the character of teacher burnout. In fact, the waves of reform had two key impacts on teacher burnout. First, they have exacerbated burnout for all groups of teachers. Second, they have had especially negative effects on certain categories of teachers, identifiable either by years of experience, race, or gender. In the next section I present data based on surveys of teachers collected during the midpoints of the different waves of school reform, as well as from a time period prior to reforms. This pre-reform data serve as an estimation of baseline levels of teacher burnout that might be found in urban school districts. It must, of course, be recognized that cohort effects would prohibit concluding the teachers surveyed in each time period are exactly comparable in their level of enthusiasm about teaching or in the challenges they faced with their students. Further, since the data are cross-sectional in nature, it would be incorrect to see the results as displaying trajectories of burnout.

BURNOUT PATTERNS IN DURING THE DIFFERENT WAVES OF SCHOOL REFORM

The data for this paper come from teachers in a single Texas school district, the largest in the state and the seventh largest in the United States, which employs in excess of 12,000 teachers and serves more than 210,000 students per year. The same ten-item burnout scale was administered to the teachers in each panel, but scores on the Teacher Burnout Scale (Dworkin 1987, 1997, forthcoming) were derived from factor scores based on a common metric across panels. This permits some limited direct comparisons of differences in burnout level between panels and across years of experience within panels. The four panels reflect different stages in the school reform efforts in the particular State and across the United States. Comparisons across panels provide indications of the changing impact of school reform activities on teacher burnout. Figure 1 displays the effects of the different school reforms on the burnout levels of different cohorts of teachers, each indicated by years of teaching experience. The independent variable is years of teaching, collapsed into categories from zero to over 30 years. The purpose for indicating year three is that it is normally during that year that teachers are reviewed for the issuance of a continuing contract (that is, the equivalent of tenure). The dependent variable is the mean burnout score for the teachers based upon a factor score across all respondents in all waves. As a factor score the measure has a mean of zero and a standard deviation of one. Positive values indicate higher levels of burnout and negative values indicate lower levels of burnout.

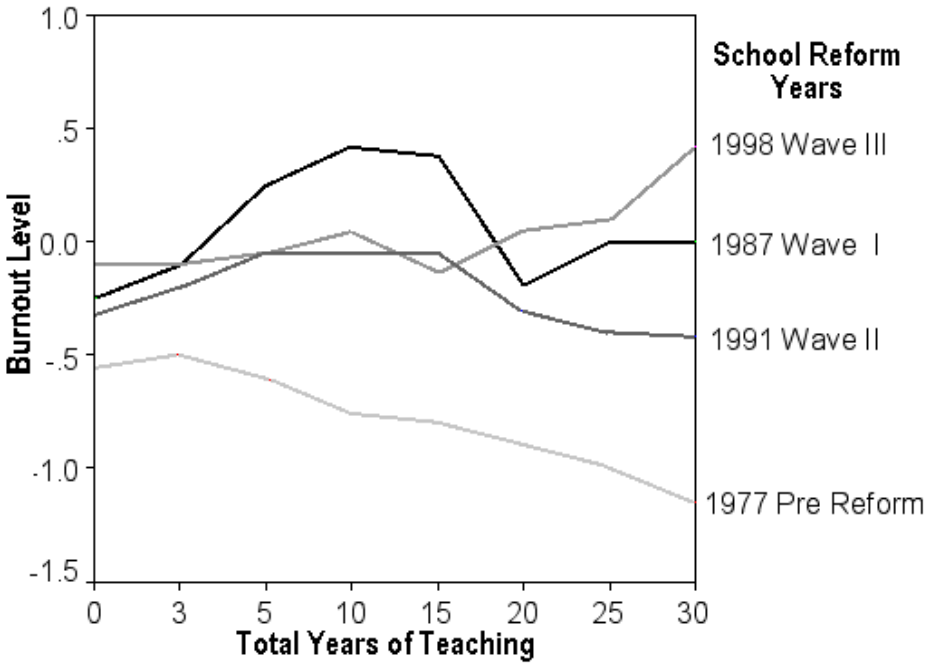


Figure 1. Burnout levels among urban teachers by years teaching

Sample sizes differed by waves: the pre-reform sample consisted of 3,165 teachers; the first wave data set had 769 teachers; the second wave data set has 246 teachers; while the third wave data was based on 727 teachers. The first wave of data, collected in 1977, reflects patterns of teacher burnout associated with different years of teaching experience in an era prior to school reform. The second-wave data, collected in 1986, depicts burnout patterns during an era following the publication of *A Nation at Risk* (1983), which the government, business, and the public questioned the extent to which schools were producing future workers who could help to maintain American economic dominance and in which centralized reforms and teacher competency testing was becoming fashionable. The 1991 data set follows the abandonment of centralization and the institution of site-based (local) decision-making at each school during the second wave. Finally, the

1998 data set reflect the most recent phase in reform, the third wave, that is, the implementation of high-stakes testing, in which the performance of students on the state-mandated standardized test has consequences for student promotion and graduation, the ratings and accreditation of schools and school districts, and the continued employment of teachers at any school.

Pre-Reform Burnout Pattern

Teachers in the different cohorts generally displayed low levels of burnout. Highest levels existed for new teachers, with a slightly higher level at year three than at the inception of teaching, probably due to the pending "tenure" decision. However, each successive cohort displayed less burnout. In the pre-reform era burnout was often the malady of the neophyte. More experienced teachers had learned to cope with the stressors associated with the teaching job. In addition, burnout tended to be higher among white teachers, as well as individuals who felt that luck, chance, and fate controlled their destinies. Likewise, teachers who were racially isolated at their schools and teachers who did not care to be assigned to schools with the racial composition found at their schools were candidates for burnout. Finally, burnout levels were particularly high among teachers whose principals defined them as expendable employees, rather than as valued colleagues (see Dworkin 1987, 1997) for more details about the characteristics of burned-out teachers in this sample).

Burnout Patterns During Wave One Reforms

The Top-Down Reform under House Bill 72, with its competency testing of teachers, changed everything. First, as can be seen in Figure 1, burnout levels are significantly higher for each cohort, but they are especially high for cohorts with between five and 15 years experience. In fact, burnout levels are three times higher for teachers with ten to 15 years experience than they had been for teachers with the same level of experience in 1977. Even new teachers displayed more burnout than did new teachers the decade earlier. However, the cause of the high burnout levels among the experienced teachers may well have been due to the legislation. By mandating competency testing of experienced teachers and imposing a career ladder in which all who passed were placed on the same rung, the legislation denied seniority and the status honor that came with experience in the classroom. Teachers with five, ten, or 15 years experience usually assumed that they were master teachers; the legislated testing and the career ladder rejected that assumption. The effect was demoralizing. Minority teachers, particularly African Americans and Hispanics, were more likely than whites to fail the written tests and were thus most likely to experience burnout (Dworkin 1997). In essence, the racial makeup of the burned-out population had changed.

Burnout Patterns During Wave Two Reforms

The 1991 data, during the implementation of site-based decision making, revealed that burnout levels had relaxed somewhat from those found in Wave One. Burnout among neophytes was higher than in the 1977 data set, but slightly lower than in 1986. However, for teachers with between five and 15 years of experience it reached a plateau, then was lower for each group with more than 15 years experience. At no point did the level reflect the patterns found prior to reform. Several things are operative here. Site-based reforms mean that decision-making may be shared, which means that teachers now have the responsibility for outcomes. When decision-making was centralized and in the hands of district administrators, teachers could always fall back on the contention that they were only following orders. Now, they had to share responsibilities, too. However, there were now interest groups competing for power: the principal, the teachers, the parents, and the local community stakeholders. In short, job stress could be exacerbated by "turf

battles," conflicts over "whose school is this, anyway?" (Dworkin and Townsend 1994). In many instances it is the white teachers in minority school that display the highest levels of burnout.

Burnout Patterns During Wave Three Reforms

The characteristic of this wave of reform is that teachers can lose their jobs if student achievement at their school is not improved. In inner-city schools the ability to raise achievement is partly due to the leadership style of the principal and partly due to skills of the teachers. In general, most teachers resemble those described by Orfield (1975) when he compared urban and suburban schools. In schools in the suburbs children come to school with so many educational resources provided by relative family affluence that teachers do not have to have much skill in order to enable their students to pass tests. In the inner-city, however, where children bring few resources from home, real teaching skill is needed and teachers soon discover that they really do not know how the teach. Older, more experienced teachers, tend to have expertise in classroom management, a skill that is necessary, but not sufficient to raise test scores. Experienced, minority teachers often have the most trouble raising scores. In addition, Texas has a system for retirement benefits called "the eighty system". That is, the combined total of one's age plus years of teaching must equal 80 in order for a teacher to retire at full benefits. Teachers with 20 to 30 years of experience are at risk. They are hoping to be able to hold on for a few more years and not be terminated because their school was too low performing or because their students' scores were too low. It may well be that the exigencies of the accountability system has led to the high burnout levels among the most experienced teachers. The highest burnout rates are found among the most experienced minority teachers.

INTRA-WAVE COMPARISONS

While the patterns of each wave of reform are different as seen in Figure 1, is it also the case that there are statistically significant differences for each experience grouping within a wave? Table 1 provides tests of significance for cohort groupings of the teachers during each wave of reform. In order to maintain adequate sample sizes categories are collapsed into four groups based on years teaching: 0-5 years, 6-10 years, 11-15 years, and 16 or more years. In the Pre-Reform sample groups with ten or fewer years teaching displayed significantly more burnout than those with more than ten years teaching. Burnout was most often the malady of the less experienced. However, during Wave One, burnout levels were significantly higher among teachers with six to 15 years experience than among other groupings. In both Wave Two and Wave Three Reforms there were not statistically significant differences among the groupings of teachers, suggesting the overall prevalence of burnout.

SUMMARY AND CONCLUSIONS

This paper has attempted to explore the relationship between school reform and teacher burnout. As a response to job stress and related to a sense of meaninglessness and powerlessness, burnout is a malady of human service professionals who are denied professional autonomy, status, and respect. In response to the appearance that American public schools are failing and that the country could be in jeopardy of losing its economic dominance, government and business have joined together to implement a series of far-reaching reforms. However, there is an apparent capriciousness to the rapidity of the reforms. Between the First Wave reforms following *A Nation at Risk* (1983) and the move toward high-stakes testing of the Third Wave reform merely a decade elapsed. Interviews with teachers conducted under the aegis of a study of factors that affected the student achievement of children in high-poverty schools (Dworkin et al. 1998) revealed that many teachers would have been more stressed and likely more burned out if it were not the case that

many had learned to cope with reforms by ignoring them. Elsewhere I have reported that the relationship between stress and burnout has become recursive. The greater the level of stress the greater the level of burnout. However, once burnout has reached a high level the case holds that the greater the burnout the less the stress (Dworkin 1997). In essence, burnout becomes a coping mechanism though which teachers cease to care and thereby experience diminished stress.

Some might ask whether it is the same groups of teachers who are burned out across the waves. That is, is it possible that the teachers who experienced higher burnout levels in 1977 as new teachers were the same people who reported high levels of burnout in 1986 as teachers with a decade of experience? It is reasonable to assume within the context of studies of a single, large, urban district that some of the teachers in one study were part of the sample in a subsequent study. However, this is not the case, as the demographics of the teachers with the high levels of burnout changed across waves. Burnout was highest among white neophytes in 1977, but highest among experienced minorities in 1986, and highest among whites in 1991, and highest of minorities in 1998. That burnout levels could change for groups suggests one other possibility, albeit not testable on data sets where identities of teachers are unknown. There is at least the possibility, if not likelihood, that burnout is situational and contextual, as is the case for commitment and other work attitude variables. In other words, burnout is likely not a permanent condition. As working conditions change or as individuals develop either coping skills or find supportive work environments, burnout may dissipate. Elsewhere I have reported that burnout does not lead to actual quitting behaviour. Perhaps it is not simply that the burned out do remain in teaching, disliking their jobs, but that they overcome burnout and learn to adapt even to frequent school reforms.

Table 1. Burnout Levels among Cohorts by Reform Waves

Pre-Reform Sample				
Cohort	**n**	**Mean**	**S.D.**	**Sig. Diff. Group**
0-5 yrs.	1,216	0.17	0.88	
6-10 yrs.	635	0.07	0.90	Groups 1,2 > 3,4
11-15 yrs.	479	- 0.12	0.90	
16+ yrs.	835	- 0.23	0.89	
Wave I Sample (Top-Down Reform)				
Cohort	**n**	**Mean**	**S.D.**	**Sig. Diff. Group**
0-5 yrs.	224	- 0.06	0.99	
6-10 yrs.	162	0.47	0.95	Groups 2,3 > 1,4
11-15 yrs.	144	0.39	0.90	
16+ yrs.	239	- 0.13	0.98	
Wave II Sample (Site-Based Reform)				
Cohort	**n**	**Mean**	**S.D.**	**Sig. Diff. Group**
0-5 yrs.	69	- 0.02	0.97	
6-10 yrs.	49	0.24	1.00	No sig. Difference among groups
11-15 yrs.	43	- 0.11	1.04	
16+ yrs.	85	- 0.07	1.02	
Wave III Sample (High-Stakes Testing)				
Cohort	**n**	**Mean**	**S.D.**	**Sig. Diff. Group**
0-5 yrs.	196	- 0.11	0.94	
6-10 yrs.	168	0.13	0.93	No sig. Difference among groups
11-15 yrs.	122	0.04	0.94	
16+ yrs.	241	- 0.05	0.97	

REFERENCES

Cedoline, A.J. (1982). *Job Burnout in Public Education: Symptoms, Causes, and Survival Skills*. NY: Teachers College Press.

Cherniss, C. (1980). *Professional Burnout in Human Service Organizations*. New York: Praeger.

Cherniss, C. (1992). Long-term consequences of burnout: An exploratory study. *Journal of Organizational Behaviour*, 13, 1-11.

Dworkin, A.G. (1987). *Teacher Burnout in the Public Schools: Structural Causes and Consequences for Children*. Albany, NY: State University of New York Press.

Dworkin, A.G. (1997). Coping with reform: The intermix of teacher morale, teacher burnout, and teacher accountability. In B. J. Biddle, T.L. Good, and I.F. Goodson. (Eds) *International Handbook of Teachers and Teaching*. London: Kluwer Academic Publishers, 459-498.

Dworkin, A.G. (forthcoming). A Scale to Assess Teacher Burnout. In L.K. Bishop and P.E. Lester (eds.). *Handbook on Tests and Measurement in Education and the Social Sciences, 2nd Edition*. Lancaster, PA: Technomic Publishing Co.

Dworkin, A.G. and Townsend, M. (1994). Teacher burnout in the face of reform: Some caveats in breaking the mold. In B. A. Jones and K. M. Borman (Eds) *Investing in United States Schools: Directions for Educational Policy*. Norwood, NJ: Ablex, 68-86.

Freudenberger, H.J. (1974). Staff burn-out. *Journal of Social Issues*, 30, 159-165.

Gold, Y. and Roth, R.A. (1993). *Teachers Managing Stress and Professional Burnout: The Professional Health Solution*. London: Falmer Press.

LeCompte, M.D. and Dworkin, A.G. (1991). *Giving Up on School: Student Dropouts and Teacher Burnouts*. Newbury Park, CA: Corwin Press.

Maslach, C. (1978a). The client role in staff burnout. *Journal of Social Issues*, 34, 111-124.

Maslach, C. (1978b). Job burnout: How people cope. *Public Welfare*, 36, 56-58.

Maslach, C. and Jackson, S.E. (1981). The measurement of experienced burnout. *The Journal of Occupational Behaviour*, 2, 99-113.

Maslach, C. (1993). Burnout: a multidimensional perspective. In W.B. Schaufeli, C. Maslach, and T. Marek (Eds.) *Professional Burnout: Recent Developments in Theory and Research*. Washington, DC: Taylor and Francis, 19-32.

National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C.: United States Government Printing Office.

Pines, A. (1993). Burnout: Existential perspectives. In W.B. Schaufeli, C. Maslach, and T. Marek (Eds.) *Professional Burnout: Recent Developments in Theory and Research*. Washington, DC: Taylor and Francis, 33-52.

Pines, A., Aronson, E. and Kafry, D. (1981). *Burnout: From Tedium to Personal Growth*. New York: Free Press.

Pines, A. and Aronson, E. (1988). *Career Burnout: Causes and Cures*. New York: Free Press.

Seeman, M. (1959). On the meaning of alienation. *American Sociological Review*, 24, 783-791.

Seeman, M. (1975). Alienation studies. *Annual Review of Sociology*, 1, 91-123.

Shinn, M. (1982). Methodological issues: Evaluating and using information. In W.S. Paine (Ed.) *Job Stress and Burnout: Research, Theory, and Intervention Perspectives*. CA: Sage, 61-82.

Sparks, D. and Hammond, J. (1981). *Managing Teacher Stress and Burnout*. Washington, DC: ERIC Clearinghouse on Teacher Education.

Swick, K.J. and Hanley, P.E. (1983). *Teacher Renewal: Revitalization of Classroom Teachers*. Washington, DC: National Education Association.

Texas State Legislature. (1984). *House Bill 72*. Austin, TX.

Tuebsing, N.L. and Tubesing, D.A. (1982). The treatment of choice: Selecting stress skills to suit the individual and the situation. In W.S. Paine (Ed.) *Job Stress and Burnout: Research, Theory, and Intervention Perspectives*. Beverly Hills, CA: Sage, 155-171.

United States Department of Education (2000). *The Condition of Education 2000*. NCES, Washington, DC: United States Government Printing Office, 2000-2602.

Action Learning: A Strategy for Change[1]

Halia Silins
Flinders University School of Education *halia.silins@flinders.edu.au*

The National Staff Development Committee of the Vocational Education and Training Sector is promoting action learning as a preferred professional development strategy to support the implementation of key competencies. This paper reports on an investigation of action learning as trialed across five training areas within the Department of Employment, Training and Further Education in South Australia. Two semi-structured hour long interviews were conducted with participating staff, one at the beginning, the other at the end of the project, and two questionnaires were administered: the Stages of Concern Questionnaire and the Myers-Briggs Type Indicator. This paper focuses on the two interviews and these results are discussed in relation to the effectiveness of action learning as a change strategy that can move an organisation for learning toward becoming a learning organisation.

action learning, organisational change, staff development, TAFE sector

INTRODUCTION

Bringing about change in any organisation is a complex matter. Organisational change in general and educational change in particular has never been a linear process (Fullan, 1993). Experienced managers of change are well aware that it is impossible to control the many factors operating during the change process. Selecting appropriate change strategies is an important step in achieving successive approximations toward desired outcomes. Can a strategy for change such as action learning help an organisation for learning become a learning organisation?

A learning organisation "is one that learns continuously and transforms itself. Learning takes place in individuals, teams, the organisation, and in the communities with which the organisation interacts. Learning is a ... strategically used process, integrated with, and running parallel to work ... . The learning organisation has embedded systems to capture and share learning" (Watkins and Marsick, 1993, p. 8-9).

Senge (1990) refers to such organisational change as a 'fundamental shift of mind'. Such a shift of mind requires all the individuals in the organisation to accept the necessity to integrate learning with everyday work. Such integration is more likely to occur when individuals are organised to work in teams. Team learning is more likely to be actioned and disseminated to other individuals or teams. Teamwork engages the individuals' commitment and capacity to learn at all levels of the organisation and propels the organisation toward becoming a learning organisation.

This concept of organisational learning suggests a dynamic approach to change that incorporates the concept of continuous improvement. It relies on the resources and experiences of its members

[1]The valuable assistance of Rob Smetak in conducting and analysing the interviews reported here is gratefully acknowledged.

and acknowledges the importance of structural, normative and cognitive factors in the process of organisational change.

An organisation moves towards becoming a learning organisation when it encourages its members to participate in organisational decisions, provides time for members to review and think about their work, provides professional development opportunities to expand knowledge and skills, removes the risks in being open and introduces structures and processes that make collaboration and teamwork irresistible.

ACTION LEARNING

The National Staff Development Committee (NSDC) has been promoting action learning as a staff development strategy to support the implementation of key competencies (Mayer, 1992) and bring about system change in the Vocational Education and Training Sector (VET). The implementation of a curriculum innovation such as the key competencies in the VET Sector is a major change initiative that requires people to change what they do and how they think about what they do.

The Department of Employment, Training and Further Education, South Australia (DETAFE) received funding from the NSDC to trial action learning as a support strategy for implementing the key competencies. An investigation of this trial was undertaken as one of seven research projects contributing to the larger project entitled *Teaching and Learning the Key Competencies in the Vocational Education and Training Sector: Research Support* (1996); a collaborative project between DETAFE(SA) and the Flinders Institute for the Study of Teaching (FIST) funded by the Department for Employment, Education and Training (DEET, reconstituted now as DEETYA).

Action learning (Revans, 1991) is a team-based, workplace activity that brings together people with a common problem or project to work out solutions or achieve project outcomes. The action learning group or set provides support and encouragement to try out new ways of doing things and new ways of thinking about things. Developed by Reg Revans (1971) as a staff development activity for managers in industry, it attempts to overcome resistance to new learning and the tendency to stay with the familiar methods and avoid taking personal and professional risks. Action learning provides participants with opportunities to pool their knowledge and skills, share learning tasks, review and reflect on their learning, question each others' views and ideas and learn how to work productively in a team. The process of action learning helps participants learn how to learn by dealing with real problems in the workplace.

The implementation of the eight key competencies (listed in Appendix 1) in the VET Sector has provided opportunities to trial action learning as a systems strategy that will: embed the key competencies in the curriculum; provide the staff development necessary to implement the curriculum; and, initiate the organisational change toward a work environment that enables the organisation to practise what it preaches.

This paper presents an investigation of action learning as a staff development activity supporting the implementation of the key competencies. The results and interpretations of this research reveal some of the existing complexities related to using the action learning process as an organisational change strategy.

Procedures of the Investigation

Five groups or sets with a total of 28 staff participated in trialing action learning at the Torrens Valley Institute, South Australia. These were:

- Electronics and Information Technology
- Veterinary and Applied Science
- Horticulture
- Community and Health Sciences
- Para Dental

The facilitators/leaders from each set formed a sixth group for support to debrief, plan and exchange issues.

Information about action learning and the key competencies for the larger investigation was collected using a range of methods that included observation, interviews and questionnaires. The focus of this paper is confined to the results of two hour long semi-structured interviews (Appendix 2 and 3) conducted with each member of the five sets at the beginning of the project and towards the end.

First Interview: Early Experiences of Participants

Beginning Concerns about Action Learning (Question 1)

The beginning concerns about the action learning process were:

- time constraints, attendance at set meetings and workloads (51 per cent);
- inadequate understanding of what participation in the project required (20 per cent);
- inadequate induction into the action learning process (12 per cent).

Some other concerns expressed at this stage were:

- getting active involvement in the meetings;
- promoting open communications;
- choosing the problem;
- establishing clear goals;
- keeping team focused;
- recording and reporting;
- difficulty with reaching consensus;
- restraining dominant personalities; and,
- managing conflict.

Understanding the Nature of Action Learning, its Strengths and Weaknesses (Question 2, 3 and 4)

Understanding the nature of action learning:

- mostly inadequate as expressed by, "opportunity to get together", a "discussion group" with a common aim, driven by a task "getting a job done" (39 per cent);
- deeper understanding of the staff development and team building purposes was demonstrated by explanations such as, a cycle of focusing on problems, working out solutions, questioning, reflecting, reviewing as a regular work pattern, "learning from each other" in a collaborative environment, "challenging viewpoints" and "going one step further than our normal interactions," (32 per cent). Two of the five sets had a majority of their responses in this category.

- no understanding or mistaken view demonstrated by, "like brainstorming", "like flexible delivery" (29 per cent).

The strengths of action learning that members of one set were experiencing related to the team building process: "expressing and resolving our feelings about the process"; learning to handle criticism; "being equal to one another and being valued for our input"; "getting feedback from others and bouncing ideas off people extends the learning and helps to make changes". Facilitation skills were recognised as a critical factor in promoting open communication and handling conflict.

A number of staff mentioned the time taken by action learning as a weakness of the process. One insightful comment from a member of a well-informed set was that "availability of time could be an issue to people who didn't understand what action learning is all about."

Self-Generated Action Learning Processes in Use (Question 5)

The processes most often identified across the sets were discussion, raising issues and ideas, challenging others, sharing tasks, and researching information on key competencies. One leader commented that what they did was not specific to action learning. Use of reflecting or reviewing was mentioned by 14 per cent of the total group.

Some individuals mentioned other relevant processes, expressing feelings, attentive listening, being inclusive, documentation and making information available through electronic mail, being experimental, using lateral thinking and reaching consensus rather than using majority decision.

Of the total group, 20 per cent, including two facilitators, suggested no processes and confessed they had no idea what they were.

What Helps and Hinders Your Contribution (Questions 6 and 7)

Members of sets seemed well aware of what helped and hindered their contributions at meetings. In any one set, nearly all of the following range of factors were recognised in both categories:

Table 1. What helps and hinders

Helping	Hindering
understanding action learning and/or key competencies	lack of understanding of action learning and/or key competencies
having time to participate	time constraints and external demands
having common focus for meetings	having unclear or different goals
establishing a supportive climate	unresolved conflict
using interpersonal skills	lack of interpersonal skills
finding resources	not carrying out tasks

The importance of facilitation skills to the functioning of the set was mentioned. Willingness to change was offered by one perceptive individual. Diversity of work roles in the sets was identified as both a help and a hindrance.

Satisfaction with Set Meetings (Question 8)

- Very satisfied with the way the set meetings were going (25 per cent) with over half of these responses coming from one set.
- Reasonably satisfied with the meetings (32 per cent).
- Uncertain (18 per cent).

Use of Recognised Processes in Set (Question 9)

Table 2. Action learning processes

Action Learning Processes	Not at all %	Some extent %	Reasonable extent %
clarifying goals and expectations	0	48	52
pooling of knowledge and skills	0	16	84
sharing learning tasks	4	40	56
review and reflect on my learning	8	44	48
questioning of old ideas and views	4	48	48
working as a team	0	28	72
trying out new ways of doing things	12	48	40
providing encouragement and support for change	12	32	56

The two processes recognised by most individuals (84 and 72 per cent) as occurring to a reasonable extent were the "pooling of knowledge and skills" and "working as a team". "Trying out new ways of doing things" and "providing encouragement and support for change" were the most variable, each with 12 per cent of members indicating no occurrence.

Do Sets Help Implementation of the Key Competencies (Question 10)

- Yes (68 per cent).
- No (32 per cent).

Ratings of Personal and Group Commitment to Implementing the Key Competencies (Questions 11, 12 and 13)

Personal and group commitment was high for nearly half the group (48 per cent).

Discrepancies between personal and group ratings indicated lesser commitment (52 per cent) with the majority of these responses coming from three sets. The highest differentials between the ratings were due to strong dissatisfaction with members who were seen as using the process for staff development rather than problem solving, frustration with lack of time, ineffective group process, unclear goals and diversity of styles and roles of members. Only one person revealed a fundamental difference with the implementation of the key competencies indicating that they should be addressed in secondary education.

Action Learning Outcomes - Second Interview

Of the 28 staff commencing the project, 22 were available for a second interview. The contracts of some of the staff involved had not been renewed in the second half of the project and two staff withdrew from the activity after the first few meetings. One set (Horticulture) was completely disbanded after their sixth meeting when most of the members' contracts were not renewed.

Resultant Changes to Curriculum in Relation to the Key Competencies (Question 1)

Three of the five original sets involved in this project had completed or were near completion of the goals set in their action plans. A number of members commented that this project raised their awareness and understanding of the key competencies and led them to consider ways of making them more explicit in their delivery.

Current Concerns Related to the Implementation of the Key Competencies (Question 2)

The concerns raised by staff were:

- Finding the time and resources to implement the key competencies effectively
- Integrating the key competencies formally into all curriculum documents
- Changing actual teaching methods and delivery
- Conveying the value of the key competencies to task-oriented practical students
- Developing appropriate methods of assessment and graded assessment
- Collaboration with industry to get consensus on the importance of key competencies
- Need for extensive staff development within sector
- Re-education of those outside, schools, industry

These concerns indicated a significant shift from the staff concerns expressed in the first interview. The beginning concerns were predominantly related to time constraints, inadequate understanding of what was required, lack of information about the process and concerns about productivity and involvement at meetings. The second interview indicated that concerns had progressed to the management of change and the impact of the innovation. Concerns with the consequences of the key competencies on students were emerging with the focus on delivery, assessment and changing the attitudes and values of students.

Some staff were overwhelmed by the demands of this change on their time and their ability to deliver the outcomes required of them. They argued that teaching key competencies should not be the responsibility of the sector but be part of the prior learning in schools. Others coped by assuming that they would not need to change their practices because students would learn the key competencies 'naturally' anyway.

Reports on the Experience of Working in the Set (Question 3 and 4)

Nearly all of the participants made some positive comments about their experience of working in the sets. Members of three of the sets were uniformly positive about the action learning process and their learning about key competencies. All members of these sets made general comments that indicated they had established supportive climates which contributed to the development of their team. The non-teaching members of one set commented that they could contribute little to the mapping task carried out by a couple of members in their set, but their understanding of key competencies had dramatically increased and, as administrators, they had gained valuable insights into a central activity of their organisation ie. teaching. Two sets attributed their success to the facilitation skills of the leaders.

The heterogeneity and variable attendance in a fourth set produced the widest range of comments about their experience of working in the set. The leader of this set had significant difficulty in focusing members' efforts. Just over half of the members in this set appeared to find value in their set's activities, whether it was involvement in one area's mapping task, an improved understanding of the key competencies or an opportunity to interact across teaching areas. Some members attributed their dissatisfaction to the complexities of facilitating such a diverse set. Other members were very task oriented and believed this process was an inefficient use of time for the outcomes gained.

Ratings of Personal and Group Commitment to Implementing the Key Competencies (Question 5, 6 and 7)

Nearly 60 per cent of the remaining participants indicated an equally high degree (between 8-10) of commitment to the implementation of the key competencies for the group and themselves. The remaining 40 percent indicated a discrepancy between their own and their set's commitment, with the majority of these, and the lowest ratings, coming from one set.

Comparing the ratings in the first interview with the second, all but one of the increased ratings occurred in three sets. The members from the fourth set have indicated the highest variation in ratings and half of the total of decreased ratings are assigned by this set. The explanations offered for low ratings and the discrepancies were: wide variety of work teams within the set; irregular attendance; just beginning to come to grips with key competencies; agreed upon task not a priority for others; not enough achieved in the time; due to administration not teaching responsibilities; too much is happening in DETAFE; and, key competencies should be developed mostly in schools.

Comments on Action Learning Processes and their Success (Question 8 and 9)

Choosing the set's focus for implementing the key competencies

Three of the four sets found this process straight forward on the whole. All except one member of the fourth set indicated that this was a long, drawn out process with no-one willing to take the initiative and marked by confusion and misunderstandings. In the end, the experience of one member's area was brought to bear on the early mapping activity of another.

Increasing understanding of the key competencies

There were members in each set (7 out of the total of 22) who stated they were already familiar with the key competencies and the activities of the set did not increase their knowledge significantly. The remaining individuals (15) indicated that they had all gained in understanding. It appears that for those staff who knew little about key competencies, the sets worked well to increase their understanding.

Raising awareness of the action learning processes

All members of one set asserted that their awareness of the action learning processes had been raised in the set. Ten members across three sets did not enhance their present understanding of action learning through the set activities. A frequent comment was that the set meetings were like their own experience of groups and therefore nothing new was offered under the name of action learning. The majority of these responses (7) came from the set in which the leader doubted that members had done anything more than come together to complete a task.

Maintaining commitment to the project throughout the set meetings

There was a strong and sustained sense of commitment in the responses of two sets. Set meeting times were adhered to and adjustments were made to agendas, or, in one set, meetings taped for the benefit of the member who could not attend. The remaining two sets recorded a great deal of variability in attendance and commitment. In one set the majority expressed difficulty with commitment due to pressures of work load and time constraints, lack of clarity and direction, frustration with process, and time wastage. Considerable commitment was required of hourly paid (PTI) staff who attended in their own time.

Level of support and resources from management

Nearly all participants perceived the support and resources provided by management as adequate to very good. Some members would have liked a more visible management presence in the project.

Strong comments surfaced with regard to the use of funds provided for PTI backfill. Most participants in this project had difficulty taking full advantage of these resources. For some, the hassle of organising a PTI backfill in an already tight schedule, or the unavailability of suitable replacements prevented gaining time relief. Sessional replacement for some staff was not possible because of the nature of their program. Set meetings during teaching break were taken out of preparation time. Non-teaching staff and sessional staff were not aware that they could tap into these funds.

Training of members in the action learning processes

Most comments were negative regarding the preparation of participants. Some members did not, or could not, attend induction workshops. Some obtained a resource folder on action learning but did not have time to work through it. Those who attended generally found the induction inadequate and confusing with no skills training.

Training of facilitator

Members of sets recognised the importance of facilitator skills and the need for induction workshops for facilitators. One set rotated leadership because of the inexperience of the assigned facilitator who provided a supportive administrative role. Members of another set were sympathetic to the difficulties that a leader would experience in facilitating a large and diverse group with varying attendance and commitment. Sufficient training and time allocation to accommodate the training was recognised as critical. Only one of the facilitators attended the induction program, the others would have preferred to have been better prepared.

Time allocation for set meetings and preparation

Two of the sets had no major problems with maintaining regular attendance. One set found itself continually renegotiating times to accommodate its members. Another set fixed meeting times but attendance was an issue. All members of the sets found difficulty allocating time for preparation work for the set meetings.

Formulating an action plan and outcomes

The amount of difficulty experienced with formulating an action plan and outcomes varied across sets. Two sets achieved this relatively painlessly, although the level of detail and consensus were mentioned as issues. Members of another set acknowledged that their plan and outcomes were mainly the work of one member's efforts and some concern was expressed about the unequal load distribution. A fourth set found the process arduous and lacking in clarity and direction. Two members were allocated the task and completed the plan independently having failed to negotiate a mutual meeting time.

Implementation of the outcomes

Members of three of the sets were confident that implementation of their outcomes would follow; two of the sets were waiting for their proposals to be printed, another set would complete their goals within two weeks and have made a commitment to follow through on implementation with or without funding. Members of the fourth set provided mixed responses to implementation

partly because most of the set are not in the area that will be implementing the outcomes. One member indicated that the documentation was still being prepared, the remainder of the set including the facilitator were uncertain about the progress of implementation at this stage.

How successful were set meetings in developing group cohesion, resolving conflicts and solving problems? (Question 10)

From members comments across the sets, three of the sets had been fairly successful in developing group cohesion. Only one set had difficulties in this area and members associated this with attendance problems, diverse work teams, changing group size and composition and personal styles.

Two sets acknowledged some interpersonal conflict which members and facilitators resolved as they occurred. All members of one set stated categorically that there were no interpersonal conflicts. Another set's members acknowledged disagreements, strong opinions, expressed frustration's but no conflict!

Members of three of the sets indicated problems were solved as they went along. The remaining set members gave a range of contrasting responses; problems were discussed but not necessarily solved, problems definitely resolved, problems not that well solved, problems explained and explored, problems solved via discussion, did not get to the stage of solving problems, fairly successfully solved because people wanted to and people gave ground.

How openly and honestly have you been able to express your views? (Question 11)

Only one member of the total group indicated that they were somewhat reserved because of lack of knowledge and experience.

Level of learning, perceived success and satisfaction (Question 12)

Only three of the participating staff indicated little learning in their set related to the action learning process. The rest of the group indicated moderate to high level of learning, with one member abstaining. The perceived success of the sets were generally rated successful to highly successful with only two members indicating a perceived lack of success. Most members indicated they were satisfied or very satisfied with their set, except for three members who indicated they were dissatisfied. The lower ratings in the three areas of learning, success and satisfaction were confined to one set. Members of this set gave group size, timing of project, lack of prior knowledge of key competencies and poor success of the action learning process as reasons for the lower ratings.

Response of other members and impact on other members of work team to set's activity (Question 13 and 14)

Two sets indicated that they took pains to keep other members of their work teams well informed. Nevertheless responses of work team members not involved in the set activities have varied from very interested to apathetic. Members of the other two sets have not reported to their own work teams yet and no formal process has been set up for disseminating information.

The impact of these activities on work teams outside the set was mostly in terms of awareness raising about the key competencies or no impact at all. Two specific examples of concrete impact was on the industry sectors. Two of the sets initiated consultation with sectors of the relevant industry with regard to the key competencies. One of these sets has transported the action learning process to the management of their work team.

Action learning as a staff development strategy (Question 15)

Members of two sets gave a resounding yes to the question of advocating action learning as a staff development strategy. Suggestions for improvement of the process were offered and these provide valuable guidelines to those managing the process.

Future implementation of action learning should consider:

- adhering more closely to the guidelines for initiating action learning, providing funds is not enough;
- careful selection and preparation of facilitators who are crucial to the success of the process;
- making clear to staff that being involved in action learning teams is a priority;
- allocating or funding time for the activity, builds commitment to the process and to the goals;
- recognising that not all staff are positively predisposed toward change and an action learning methodology;
- the effectiveness of an action learning set is influenced by the personalities of the members as well as the task;
- working on program issues in sets develops ownership of initiated changes;
- thoroughly informing participants of the process to obtain voluntary commitment.

Enthusiasm for action learning parallels members' perceived success of their set's activities. Members who experienced less satisfaction with their sets were more equivocal about action learning.

Some of the comments qualifying the use of action learning were, only if:

- management conducts thorough preparation, induction and sets clear expectations of what it hopes to achieve;
- members have some interpersonal skills with conflict resolution skills;
- the sets have strong facilitators and a purpose;
- participants' awareness is raised about what is involved in action learning before volunteering and beyond the provision of reading materials (which staff do not read);
- small groups are used and extended periods of confusion reduced;
- members of sets commit to the identified goals;
- management addresses the issue of funding staff development activities for temporary staff.

What issues did facilitators experience? (Question 16)

Facilitators experienced the following problems:

- members not well informed about the demands of being involved in this project;
- inadequate induction;
- not enough lead time for people to commit to this staff development process and cover other work responsibilities;
- end of the year inappropriate timing for project;

- not to have too large a group;
- maintaining attendance;
- keeping people on task yet not directing them;
- keeping a balance between task and group processes;
- time for recording and documenting process; and,
- the constraints of time and work load demands interfering with the process.

Action Learning as a Staff Development Activity

The effectiveness of action learning as a change strategy is dependent on its effectiveness as a staff development activity.

Effective approaches to staff development recognise that:

- change is a process and that bringing about behavioural and attitudinal change takes time and effort;
- successful staff development programs respond to the assessed needs of the participants and match the content and activities to their developmental level;
- applying the principle of self-management fosters an involvement in the learning and a commitment to change;
- participants need to be involved in decisions about program content and activities and the processes for achieving the objectives;
- effective staff development acknowledges and capitalises on the experience and existing expertise of the participants;
- it is important to take account of participants' changing concerns at different stages in the process;
- the purpose of innovations should be made clear and related to the achievement of the overall organisational goals and objectives;
- those managing the process must provide opportunities to implement new learnings and provide ongoing support and feedback on progress.

Effectiveness of staff development programs will be reduced when:

- activities are imposed and piecemeal;
- activities are not work-based and there are few opportunities to practise and implement new learnings;
- managers do not demonstrate commitment nor provide tangible support; and,
- the culture of the workplace fails to support program goals.

Most of the participants in this project indicated that they were not well enough informed about what was involved to have made a voluntary commitment to the project. Some felt that they had been misled about the extent of the demands of the project. For those that attended, preparatory workshops were confusing and presented an oversimplified picture of what was required.

Action learning is supposed to draw people together with a common work problem to find a solution or complete a task. The individuals in this study did not initiate the process. The availability of funds was the catalyst for their formation. The sets were brought together to trial action learning by addressing the issues around the implementation of key competencies. The dual

nature of the sets' activities was confusing and resulted in negativity toward the action learning process which was not well understood.

Some sets had difficulty in identifying clear goals and achievable outcomes because of the diversity of roles and interests in their set. The leader support meetings indicated that goals were not always reached by consensus and one or two members of one set were left to achieve an outcome because the goals were not necessarily of relevance to all members.

Action learning provided opportunities for staff to learn from each other and share resources. Those who took the opportunity reported an enhanced awareness and understanding of both the key competencies and action learning. Pooling of knowledge and skills was seen as a process being used by most of the individuals.

To some extent, participants experienced support and encouragement to try new things through the set meetings. However, members in some sets indicated it was not occurring at all. Some participants indicated that as a result of the action learning activity they were planning to introduce new methods of delivery and new systems to their teaching. However, these examples were the exception rather than the rule. Learning new skills requires experimentation and practice. Implementing the key competencies seemed premature for most members.

The evidence from the sets suggests that at least in two of the sets the skills of working in teams were being established. Team building requires an appreciation of people's differences and strengths. Set members found difficulty with working in diverse groups. Some individuals reported diversity as a help to the set and others a hindrance. The sets that appeared more effective were generally homogeneous. The tendency was to emphasise the task in some sets so that interacting with a diverse group became a hindrance if interpersonal learnings were not valued.

A number of members from two sets reported no conflict despite the evidence from interviews and leader meetings to the contrary. Conflicts are an inevitable product of engagement. Avoiding conflicts retards the individual's and group's development. Effective facilitation skills are required to ensure that conflicts are resolved and that the process remains constructive.

Participants in this project were satisfied with the funding support provided for this project in principle. In fact, however, most found it difficult to draw on this support for a variety of valid reasons. Regular meetings of set leaders with the project manager provided some feedback and support. Leaders had the opportunity to learn from each other how to conduct set meetings and facilitate the process.

Lewin (1951) maintained that learning is facilitated in an organisation by establishing an environment of tension between concrete experience and analytic detachment (Rait, 1995). When individuals in an organisation examine their own practices using data from actual workplace events and test new organisational actions, shifts in practice are more likely to occur. Bringing about change in the workplace involves each of the steps of the experiential cycle - thinking, doing, evaluating and reflecting (Kolb, 1984). To promote staff development and change, facilitators need to help individuals and teams move through this cycle. Multiple cycles of action and reflection result in collaborative inquiry (Dewey, 1938).

Working in teams can lead to reflection and questioning of old ways of doing things. Exchanges about work practices in a supportive group raises awareness of what we do and why we do it. Some examination of work practices occurred in the sets, but deeper inquiry about work practices was uncommon. The evidence suggests that members from two sets were consciously moving through this cycle in order to maximise learning.

Change in staff's attitudes and beliefs can occur as a consequence of a change in behaviour. Interacting with others can highlight contradictions that may exist between practice and underlying beliefs. Participants did indicate that their attitude towards action learning changed at the end of the project and became more positive.

Action Learning as a Strategy for Change

Action learning is hard work, often frustrating and sometimes threatening because individuals are expected to practice more open communication than their every day activities require. They have to become interdependent learners when they have habituated to independent learning. The ultimate aim of teamwork of this kind is organisational and social change. System change requires behaviour and attitude change which individuals may resist (NSDC Commissioned Project, 1993). Action learning requires considerable investment of energy and time to be a successful strategy for bringing about change.

Implementation of the key competencies into the VET Sector is a significant change initiative. It requires substantial professional development support to ensure a shared understanding of what key competencies mean across the sector, develop staff expertise in the eight key competency areas, provide the opportunity, knowledge and skills required to incorporate the key competencies into courses, and, ensure the key competencies are appropriately embedded into the industry competencies.

A significant factor is that the key competencies that action learning methodology is being used to advance in the VET sector, are the competencies that are required to be put into practice in action learning. Both action learning and the expectation that VET Sector staff will teach the key competencies successfully, presuppose a high level of attainment of the key competencies by the staff involved.

In this study, staff had very high levels of personal concerns around participating in the action learning project. Elevated personal concerns occur when individuals face the uncertainty and ambiguity that is associated with change. Such personal concerns create significant resistance and reluctance toward implementing change and need to be reduced. Individuals who are used to working independently find teamwork personally threatening, time consuming and frustrating. Teamwork develops a learning orientation, flexibility and openness to new ideas and new ways of doing things that remain undeveloped when working in isolation.

If collaborative teams become the basic unit of work preparation and performance management, then they can provide:

- an ongoing context and support for action learning processes;
- a way of managing the linking of curriculum design, delivery and assessment;
- a focus for continuous, work-based staff development; and,
- opportunities for the implementation of innovations such as key competencies.

The increasingly part-time nature of DETAFE staff inhibits this kind of teamwork even though it facilitates the management of outcomes and quality control. If teamwork is seen as work-integrated staff development, then its importance needs to be recognised as a required part of teaching preparation for performance enhancement and as a quality assurance activity.

Most of the individuals in this study understood action learning to be nothing more than working together in a group to solve problems. The implicit belief of these individuals was that teamwork was not the most effective or efficient way "to do things around here", particularly under the

present work demands and time constraints. Some individuals expressed a belief in the value of teamwork that was not reflected in their preferred practice.

This study demonstrates that the success of action learning as a change strategy depends on a number of factors. These are:

- preparation of the staff to participate in the project;
- selection and preparation of the facilitators;
- management support, funding and feedback;
- time constraints, workloads and attendance;
- needs and concerns of all the participants in the project;
- setting clear goals around a common problem;
- making team building and staff development a priority; and
- employing the thinking, doing, evaluating, and reflecting cycle.

CONCLUSION

As a change and staff development strategy, action learning requires an organisational culture of collegiality and cooperation. DETAFE will need to enunciate these principles in their organisational objectives, model the processes in their own work practices and provide resources in the form of release time and support. Nothing short of a cultural transformation is required in DETAFE to meet the challenge of becoming a learning organisation and improve performance in the face of present economic constraints.

Action learning is a process that requires members of an organisation to work together to solve problems through action and reflection. It has the potential to become a multi-purpose organisational activity depending on the extent to which it is integrated into the organisation's systems and made central to the organisational culture. Action learning processes promote reflection, mentoring and collaboration and cast employers into the role of continuous learners who are capable of both investigating and improving work practices.

Action learning cuts across the prevailing trend within education and training institutions to work in isolation. The group work promotes problem solving and team building as well as provides a vehicle for work-place linked ongoing professional development. Ultimately, action learning aims to change behaviours, instil new skills and revise attitudes and beliefs underlying work practices. Collaborative work-based teams of this kind have the potential to transform an organisation for learning into a learning organisation.

REFERENCES

Dewey, J. (1938) Logic: The theory of inquiry. New York: Holt, Rinehart and Winston.

Flinders Institute for the Study of Teaching (1996) *Teaching and Learning the Key Competencies in the Vocational Education and Training Sector: Research Support.* Research Report supported by a Grant under the Key Competencies program from the Commonwealth Department of Employment, Education and Training. Adelaide: Flinders University.

Fullan, M. (1993) *Change forces. Probing the depths of educational reform.* New York: The Falmer Press.

Kolb. D.A. (1984) Experiential learning: Experience as the source. Englewood Cliffs, NJ: Prentice Hall.

Lewin, K. (1951) *Field theory in social science.* New York: Harper and Row.

Mayer, E. (Chair) (1992) *Putting General Education to Work: The Key Competencies Report.* (Report of the Committee to advise the AEC and MOVEET on employment-related key competencies for post-compulsory education and training) Melbourne: The Australian Education Council and the Ministers for Vocational Education, Employment and Training.

NSDC Commissioned Project, (1993) *Action learning in Vocational Education and Training. Volume 1.* Chadstone, VIC: National Staff Development Committee.

Rait, E. (1995) Against the current. Organizational learning in schools. In S.B.Bacharach and B. Mundell (Eds.), *Images of schools. Structures and roles in organizational behavior* pp. 71-107. Thousand Oaks, California: Corwin Press.

Revans, R.W. (1971) *Developing effective managers.* London: Praeger Publishers.

Revans, R.W. (1991) Action learning: Its origins and nature. In M. Pedler (ed) *Action learning in practice* pp. 3-16. Aldershot: Gower Press.

Senge, P. (1990) *The fifth discipline: The art and practice of the learning organization.* Sydney: Random House.

Watkins, K. E. and Marsick, V. (1993) *Sculpting the learning organization.* San Francisco: Jossey Bass.

APPENDIX 1

Key Competencies

1. **Communicating Ideas and Information**
   The capacity to communicate effectively, with others using the range of spoken, written, graphic and other non-verbal means of expression.

2. **Using Mathematical Ideas and Techniques**
   The capacity to use mathematical ideas, such as number and space, and techniques, such as estimation and approximation, for practical purposes.

3. **Using Technology**
   The capacity to apply technology, combining the physical and sensory skills needed to operate equipment with the understanding of scientific and technological principles needed to explore and adapt systems.

4. **Solving Problems**
   The capacity to apply problem-solving strategies in purposeful ways, both in situations where the problem and the desired solution are clearly evident and in situations requiring critical thinking and a creative approach to achieve and outcome.

5. **Working with Others and in Teams**
   The capacity to interact effectively with other people both on a one-to-one basis and in groups, including understanding and responding to the needs of a client and working effectively as a member of a team to achieve a shared goal.

6. **Collecting, Analysing and Ordering Information**
   The capacity to locate information, sift and sort information in order to select what is required and present it in a useful way, and evaluate both the information itself and the sources and methods used to obtain it.

7. Planning and Organising Activities
The capacity to plan and organise one's own work activities, including making good use of time and resources, sorting out priorities and monitoring one's own performance. (Mayer Report, 1993, p. 3)

8. Developing Cultural Understanding
The capacity to apply an understanding of cultures when carrying out workplace tasks, including commitment to organisational goals such as quality, safety, efficiency, teamwork, security, environmental protection, customer service and personal development, and interacting with people from widely different backgrounds and cultures in the achievement of these common work goals. It involves respect for rights and responsibilities, tolerance and sensitivity and an appreciation of the wide diversity of backgrounds, languages, beliefs, customs and traditions that are spread across Australian society and its labour force.

APPENDIX 2

Semi-Structured Interview Questions

1. Action Learning concerns response (attached)
2. Describe what you understand by Action Learning
3. What do you see as the strengths of Action Learning? (or what do you like about it?)
4. What do you see as the weaknesses (dislikes) of Action Learning?
5. What Action Learning processes do you use?
6. What helps you to contribute at the set meetings?
7. What hinders your contribution at the set meetings?
8. How satisfied are you with the set meetings at this stage?

9.

Action Learning is said to provide opportunities for:	Not at all	Some extent	Reasonable extent
clarifying goals and expectations			
pooling of knowledge and skills			
sharing learning tasks			
review and reflect on my learning			
questioning of old ideas and views			
working as a team			
trying out new ways of doing things			
providing encouragement and support for change			

10. How helpful are the set meetings in your implementation of the Key Competencies?
11. How would you rate your commitment to implementing the KC's on this scale? Use "A" on the scale below.
12. How would you rate the group's commitment? Use 'B" on the scale below.

0 1 2 3 4 5 6 7 8 9 10

13. If gap exists, explain.

APPENDIX 3

Second Semi-Structured Interview Questions

1. What changes, if any, have you made to your curriculum in relation to KCs? Give examples.
2. What are your current concerns regarding the implementation of KCs?
3. Can you make some general comments regarding your experience of working in the set?

   (Rob look for positive and negative orientation to action learning)
4. How helpful were/are the set meetings in your implementation of the KCs?
5. How would you rate your commitment to implementing the KCs on this scale? Use "A" on the scale below.
6. How would you rate the group's commitment? Use "B" on the scale below.

   0 1 2 3 4 5 6 7 8 9 10
7. If gap exists, explain.
8. Here is a list of some of the processes that you've been involved in. Make some comments about any of the processes on this list.
   1. Choosing the set's focus for implementing KCs
   2. Increasing your understanding of the KCs through the set
   3. Raising your awareness of the action learning processes
   4. Maintaining your commitment to the project throughout the set meetings
   5. Level of support from management and resources
   6. Your training in the action learning processes
   7. The training of the facilitator
   8. Your time allocation for set meetings and preparation
   9. Formulating your set's action plan and outcomes
   10. Your implementation of the outcomes
   11. The attendance and participation at set meetings
9. How successful have each of these processes been?
10. How successful were/are the meetings of the set in:
    - developing group cohesion
    - resolving interpersonal conflicts
    - solving problems
11. How open and honest have you been able to be in expressing your views to the set?
12. Circle your own level of:

    learning in the set 0 1 2 3 4 5

    perceived success of the set 0 1 2 3 4 5

    satisfaction with the set 0 1 2 3 4 5
13. How have other members of your work team responded to this set's activity?
14. Have any of the set's ideas impacted on the work team outside of the set?
15. Would you advocate action learning as a staff development strategy?
16. In your role as a facilitator what issues/problems did you experience?

Change in differences between the sexes in mathematics achievement at the lower secondary school level in Australia: Over time

Tilahun Mengesha Afrassa
Department of Education, Training and Employment

John P. Keeves
Flinders University School of Education

In this paper an investigation is reported on whether changes have occurred in the differences between the sexes in mathematics achievement at the lower secondary school level over the 30 year period from 1964 to 1994. In order to make meaningful comparisons the mathematics test scores from the three studies conducted in Australia under the auspices of the International Association for Evaluation of Educational Achievement were brought to a common interval scale using Rasch measurement procedures. The scale scores are used to examine differences between boys and girls in mathematics achievement on the three occasions as well as the changes that have occurred between occasions. No significant sex differences in mathematics achievement are found on each of the occasions. However, a significant decline in mathematics achievement is recorded for boys between 1964 and 1994, but not for girls. The decline in mathematics achievement over this 30 year period for boys is equivalent to nearly one year of mathematics learning, while the drop for girls is only approximately equivalent to half a year of mathematics learning.

Gender differences, mathematical achievement, Rasch modelling

INTRODUCTION

During the past 30 years the issue of differences between the sexes in mathematics achievement in Australian schools has attracted considerable research, discussion and recommendations for change in policy and practice. Major reviews of research conducted in Australia by Leder and Forgasz (1992) and Barnes and Horne (1996) in gender and mathematics learning refer to "gender" rather than "sex" differences. However, Megarry (1984) has argued for the use of "sex" to refer to the biological category to which a person belongs, and for the use of "gender" to denote "the set of meanings, expectations and roles that a particular society ascribes to sex". In this paper interest is primarily focused on estimated differences in mathematics achievement between male and female students, without the examination of gender related influences. As a consequence the differences considered are commonly referred to here as "sex differences".

Leder (1992) has also reviewed changing perspectives in the area of gender differences in mathematics learning from an international viewpoint. However, this review does not examine factors that account for differences either across countries or over time. Articles by Keeves (1973) and Baker and Jones (1993) examine some of the data sets that are considered in this paper from a cross national perspective. Their investigations address particular relationships associated with such differences and provide evidence that the estimated differences between the sexes in mathematics achievement are at least in part, societally based. From these findings there is clear

evidence that gender based effects have an influence on the learning of mathematics. However, it is necessary to clarify whether significant differences in mathematics achievement between male and female students can be detected in Australia and at what year and age levels these are found as well as whether levels of achievement have changed over time.

Unfortunately much of the research into differences in mathematics achievement between the sexes in Australia has suffered from several serious shortcomings: (a) selection bias, since comparisons are made using groups that are not representative of a target population which is complete and has not suffered from self selection; (b) the sampling procedures employed are inadequate because, while large numbers of students are tested, they are drawn from too few schools; and (c) the estimation of error for significance testing fails to take into consideration the use of a cluster sample design. If sound comparisons are to be made for the detection of differences between the sexes in mathematics achievement, then large representative and random samples of students must be tested and significance tests with appropriate estimates of error must be employed. Unless, the research findings reported are free from these problems, any discussion of these findings and the recommendations developed from such findings are likely to be highly misleading.

The problems associated with the inappropriate use of significance tests in situations where schools are selected as the primary sampling unit and students are selected from within schools have been recognised for over 30 years. Sometimes very crude estimates of error have been calculated which sought to make allowance for the use of a cluster sample design (for example, Keeves, 1968). In general, the problem has been ignored, because appropriate computer programs have not been available for the accurate estimation of sampling and measurement errors. In the past few years computer programs have become available, for example WesVarPC (Brick et al. 1997) and HLM (Bryk, Raudenbush and Congdon 1992), that can be readily used to estimate errors that would permit the testing of such data for statistical significance. Under these circumstances, in order to examine differences between boys and girls in mathematics achievement, it would seem necessary for analyses to be undertaken with data sets that involve large random samples, with the nested structure of students within schools clearly identified, so that appropriate estimates of error can be calculated prior to testing for statistical significance.

The aim of this paper is to examine whether differences can be detected between the sexes in mathematics achievement in Australia at the lower secondary school level. Furthermore, this paper considers whether changes have occurred over time in levels of mathematics achievement as might be expected from the programs that have been introduced in Australia over the period under survey.

Several questions must be addressed before the analyses can be undertaken.

1. The samples selected must not be confounded by the effects of selection bias. This requires that the samples are drawn from schools across Australia, at a stage of schooling where the study of mathematics is compulsory and prior to the period either where some students have dropped out from school or where the study of mathematics is optional.

2. It is necessary to establish that the instruments employed to measure mathematics achievement are assessing one dimension and that the examination of a single mathematics score is meaningful, rather than separate dimensions for the different branches of mathematics at the lower secondary school level - as arithmetic, algebra and geometry.

3. Furthermore, in order to make meaningful comparisons over time, it is necessary to bring the different measures of mathematics achievement on the different occasions to a common scale.

This can be done through the use of the Rasch scaling of the scores using concurrent equating procedures, provided the items and the persons satisfy the requirement of unidimensionality, and there are common items in the tests used across the different occasions.

SAMPLING PROCEDURE

In the First International Mathematics Study (FIMS), (Keeves, 1968) conducted in 1964, two groups of students participated, 13-year-old students in Years 7, 8 and 9 and students in Year 8 of schooling. In total 2275 male and 2044 female students participated in this study. In FIMS only government schools in New South Wales (NSW), Victoria (VIC), Queensland (QLD), Western Australia (WA) and Tasmania (TAS) participated. In the Second International Mathematics Study (SIMS), (Rosier, 1980) which was administered in 1978, nongovernment schools and the Australian Capital Territory (ACT) and South Australia (SA) were also involved as well as those states that participated in FIMS. Subsequently, in the Third International Mathematics Study (TIMS), (Lokan, Ford and Greenwood, 1996) which was conducted in 1994, government and nongovernment school students in all states and territories including the Northern Territory were involved. In the TIMS study 6089 male and 6761 female students were tested at the lower secondary school level.

In 1964 and 1978 the samples were age samples and included students from Years 7, 8 and 9 in all participating states and territories, although in FIMS a Year 8 sample was also tested. In TIMS the samples were grade samples drawn from Years 7 and 8 or Years 8 and 9. In ACT, NSW, VIC and TAS Years 7 and 8 students were selected while in QLD, SA, WA and NT samples were drawn from Years 8 and 9, enabling a Year 8 Australia wide sample to be derived for purposes of comparison with the 1964 data.

Therefore, to make the most meaningful possible comparison of mathematics achievement of boys and girls over time by using the 1964, 1978 and 1994 data sets, the following steps were taken.

The 1964 students were divided into two groups 13-year-old students in one group (FIMSA) and all Year 8 students including 13-year-old students at that year level as the second group (FIMSB) since in addition to an age sample, a grade sample had also been drawn. It is important to observe that 13-year-old students in Year 8 were considered as members of both groups. In the first group, students were chosen for their age and in the second group for their year level. The 1978 students were chosen as an age sample and included students from both government and nongovernment schools. In order to make meaningful comparisons between the 1978 sample and the 1964 sample, the 1978 government school students were divided into two groups. The first group included all government school students who participated in the study (SIMSG), and the second group included all government school students in the five states excluding students from SA and ACT (SIMSR).

Meanwhile, in TIMS the students were chosen as a grade sample. The common sample for all states and territories was Year 8 students. In order to make the TIMS samples comparable with the other samples, only Year 8 government school students in the five states that participated in FIMS and SIMS are considered as the TIMSR data set in this study.

ANALYTICAL PROCEDURES EMPLOYED IN THE STUDY

In this study the procedures employed to compare the achievement differences between male and female students on the three occasions involved the use of the Rasch model to scale students' responses to the mathematics test items. The Rasch model has been shown to be the most robust of the item response models (Sontag, 1984), and was used in this study primarily to equate

students' performance in mathematics on a common scale for the Australian investigations conducted in FIMS, SIMS and TIMS.

Unidimensionality

In order to employ the Rasch model for calibrating the items in the mathematics tests it was necessary to examine whether or not the items satisfied the requirement of unidimensionality (Hambleton and Cook, 1977). If the items were found not to satisfy the condition of unidimensionality, it would not be possible to employ the Rasch procedures for the calibration of the tests. Furthermore, it would not be meaningful to compare differences between the sexes with respect to mathematics achievement since a total score would be meaningless.

Consequently, confirmatory factor analysis procedures were employed to test the unidimensionality of the mathematics test items using the LISREL computer program (Jöreskog and Sörbom, 1992).. Confirmatory factor analysis is a statistical procedure employed for investigating relations between a set of observed variables and the underlying latent variables (Byrne, 1989; Spearritt,, 1997). The results of the confirmatory factor analyses of FIMS and SIMS data sets revealed that a nested model in which the mathematics items were assigned to three specific correlated first-order factors of *Arithmetic*, *Algebra* and *Geometry* as well as a general higher order factor, which was labelled as *Mathematics* provided the best fitting model. In addition, in the confirmatory factor analyses undertaken, no evidence was found to reject the assumption of the existence of one general factor involved in the mathematics tests, in so far as in the nested model the *Mathematics* factor extracted more of the total variance than did the specific first-order factors taken together. Therefore, the mathematics test items in the FIMS and SIMS studies are considered to satisfy the requirement of unidimensionality. The item cluster-based design procedure employed in the construction of the TIMS data sets would seem to preclude the use of confirmatory factor analysis to test the unidimensionality of the TIMS data set. and confirmation of unidimensionality must be provided by the introductory steps in the Rasch analysis.

Effect of mathematics learning in one year

It is possible since the TIMS project tested at two adjacent grades to estimate the gain between the lower grade and the upper grade for the Australian sample and thus to interpret the calibrated effect size in terms of years of mathematics learning at the lower secondary school level. The difference between the lower and the upper grade levels show the growth on mathematics achievement over one year in Australian lower secondary schools. Thus to examine the growth between grade levels, the estimated mean scores difference between the lower and the upper level students was compared using the WesVarPC 2.11 computer program (Brick et al. 1997). The results of the comparisons are presented in Table 1.

The estimated mean score for the lower level was 520 centilogits, while for the upper level students, it was 557 centilogits. The difference was 37 centilogits in favour of the upper level students. Hence, the growth in achievement per year in mathematics performance in Australian lower secondary schools was 37 centilogits. The effect size and t-value were 0.30 and 3.96 respectively. Thus, the mean difference between the lower and the upper grade levels, namely 37 centilogits is equivalent to an effect size of 0.30. Keeves (1992) reported that 38 centilogits was found to be equivalent to a year of science learning between the 10 and 14 year-old levels in the Second IEA Science Study in Australia. Therefore, this information allows the differences between the achievement level of the different groups to be interpreted in terms of practical significance as well as statistical significance. Thus, the mean difference between the lower and the upper grade

levels in Australian lower secondary schools in 1994 was practically and statistically (at the 0.01 level) significant. Hence it is estimated their learning in one year between Year 7 and Year 8 is 37 centilogits and

1 centilogit = 1 week of mathematics learning,

4 centilogits = 1 month of mathematics learning.

Table 1. Comparisons between the Lower and Upper group Levels in TIMS

Levels	Mean	Standard deviation SD	Standard error SE	Sample size SS	Design effect DEFF	Effective sample size ESS	Effect size ES	t-value	Significant level Sig-L
Lower	520	118.8	7.0	5599	19.44	288.0	0.30	3.96	0.01
Upper	557	127.7	6.2	7253	17.36	417.8			

Effect size

In this paper both the standardized effect size and the magnitude of effect on the calibrated scales are used to examine the level of practical significance of the differences between FIMS, SIMS and TIMS in mathematics achievements over time. The following formula was employed to calculate an effect size value.

$$\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

Where $\overline{x}_1$ = estimated mean score for group one;
$\overline{x}_2$ = estimated mean score for group two;
s_1 = standard deviation of group one; and
s_2 = standard deviation of group two.

In this study effect size values less than 0.20 are considered as trivial, while values between 0.20 and 0.50 are considered as small. Furthermore, effect size values between 0.50 and 0.80 are taken as moderate and values above 0.80 are treated as large (Cohen, 1992).

The t -test

In order to determine the level of statistical significance between the mean scores on FIMS, SIMS and TIMS in mathematics achievement a *t*-statistic was calculated, which took into account errors from three sources: (a) sampling error, (b) errors of calibration, and (c) equating error. Since the samples all involved a cluster sample design with schools sampled with a probability proportional to size at the first stage and students sampled from within schools at the second stage, it was necessary to use the WesVarPC (Brick et al. 1997) computer program to test the data for statistical significance, taking into account both the stratification and cluster sample design employed in all three studies.

Developing a common mathematics scale

The calibration of the mathematics data permitted a scale to be constructed that extended across the three groups, namely FIMS, SIMS and TIMS students on the mathematics scale. The fixed point of the scale was set at 500 with one logit, the natural metric of the scale, being set at 100 units. The fixed point of the scale, namely 500 was taken as the mean of the difficulty level of the calibrated items in the FIMS test administered in 1964.

Rasch Analysis

Three groups of students namely FIMS (4320), SIMS (5120) and TIMS (12850) were employed in the calibration and scoring analyses. The necessary requirement for calibration in Rasch scaling is that the items and persons must fit the Rasch scale. In order to examine whether or not the items and persons fitted the scale, it was important to evaluate both the item fit statistics and the person fit statistics. The results of these analyses are presented below.

Item Fit statistics

One of the key item fit statistics is the infit mean square (INFIT MNSQ). The infit mean square measures the consistency of fit of the students to the item characteristic curve for each item with weighted consideration given to those persons close to the 0.5 probability level. The acceptable range of the infit mean square statistic for each item in this study was taken to be from 0.77 to 1.30 (Adams and Khoo, 1993). In calibration, items that do not fit the Rasch model and which are outside of the acceptable range must be deleted from the calibration analyses (Wright and Stone, 1979). Hence, in FIMS two items (Items 13 and 29), in SIMS two items (Items 21 and 29) and in TIMS one item [(Item T1b no 148) with one item (no. 94) having been excluded from the international TIMSS analysis] were removed from the calibration analyses due to the misfitting of these items to the Rasch model.

Case Estimates

A second way of investigating the fit of the Rasch scale to the data is to examine the estimates for each case. The case estimates give the performance level of each student on the total scale. In order to identify whether the cases fit the scale or not, it is important to examine the case OUTFIT mean square statistic (OUTFIT MNSQ) which measures the consistency of the fit of the persons to the student characteristic curve for each student, with special consideration given to extreme items. In this study, the general guideline used for interpreting **t** as a sign of misfit is if t> 5 (Wright and Stone, 1979, p. 169). Thus, if the OUTFIT MNSQ value for a person has a |*t – value*| greater than 5, that person does not fit the scale and is deleted from the analysis. In this analysis no person was deleted, because the |*t – value*| for all cases was less than 5. However, students with a zero score or with a perfect score were automatically excluded from the calibration procedure, since they would not provide useful information for the purposes of scale calibration, although such students were necessarily included in the scoring of the data.

COMPARISON OF SEX DIFFERENCE IN FIMS, SIMS AND TIMS

In Table 2 the mathematics achievement levels of boys and girls in FIMSA, FIMSB, SIMS and TIMS are compared (see Table 2). The paper compares differences in mathematics achievement between the two sexes on the three occasions. These comparisons for Australia in 1964 and 1978, differ from those carried out in previous studies (Keeves, 1968; Moss, 1982) in that proper account can now be taken for the complex design of the samples employed in testing for statistical significance.

However, in testing for significant differences, while multiple comparisons are involved no use is made of the Bonferroni Adjustment (Finn, 1997), because the thrust of the comparisons is more toward the detection of no differences, than towards the detection of highly significant ones.

The first comparison to be discussed is between FIMS Students.

Table 2. Descriptive statistics for mathematics achievement of all students for the three occasions by sex

FIMS	FIMSA		FIMSB		Total	
	Male	Female	Male	Female	Male	Female
Mean	462.0	457.0	455.0	448.0	459.0	453.0
Standard Deviation	101.0	90.0	87.0	76.0	98.0	87.0
Jackknife Standard Error	9.7	7.3	10.1	7.5	9.3	6.5
Design Effect	14.9	9.6	17.3	11.2	20.5	11.4
Sample Size	1530	1386	1619	1462	2275	2044

SIMS	Government		Non-government		Total		Restricted	
13year-old students	Male	Female	Male	Female	Male	Female	Male	Female
Mean	442.0	443.0	488.0	464.0	453.0	448.0	440.0	443.0
Standard Deviation	105.0	100.0	117.0	97.0	110.0	99.0	105.0	99.0
Jackknife Standard Error	4.2	5.1	14.9	9.6	5.2	4.6	4.7	5.8
Design Effect	3.4	5.0	9.5	5.4	6.2	5.2	3.2	4.8
Sample Size	2095	1894	580	551	2675	2445	1614	1424

TIMS	Government		Non-government		Total		Restricted	
Year 8 students	Male	Female	Male	Female	Male	Female	Male	Female
Mean	423.0	431.0	488.0	474.0	443.0	448.0	422.0	431.0
Standard Deviation	125.0	124.0	117.0	110.0	126.0	121.0	126.0	126.0
Jackknife Standard Error	7.7	8.4	10.2	8.4	7.0	6.4	8.4	9.2
Design Effect	9.4	9.9	8.5	9.6	11.1	10.7	8.9	9.4
Sample Size	2479	2168	1111	1633	3590	3801	2030	1755

	Mean difference	Effect size	Effect of learning in a year	t-value	Sig level
FIMSA Male vs Female	5	0.05	5 weeks	0.41	Not Sig
FIMSB Male vs Female	7	0.09	7 weeks	0.56	Not Sig
FIMS Male vs Female	6	0.06	6 weeks	0.53	Not Sig
SIMS Government Male vs Female	-1	-0.01	1 week	-0.15	Not Sig
SIMS Nongovernment Male vs Female	24	0.22	6 months	1.35	Not Sig
SIMS Male vs Female	5	0.05	5 weeks	0.72	Not Sig
SIMS Restricted Male vs Female	-3	-0.03	3 weeks	-0.40	Not Sig
TIMS Government Male vs Female	-8	-0.06	2 months	-0.70	Not Sig
TIMS Nongovernment Male vs Female	14	0.12	3.5 months	1.06	Not Sig
TIMS Male vs Female	-5	-0.04	5 weeks	-0.52	Not Sig
TIMS Restricted Male vs Female	-9	-0.07	9 weeks	-0.72	Not Sig

Restricted =Those groups of students in SIMS which are comparable with FIMSA, and students in TIMS which are comparable with FIMSB

Comparisons of Sex Differences between 1964 Australian Students

Table 2 and Figures 1 and 2 show the three comparisons which are considered in FIMS. The estimated mean score differences for the three comparisons are seemingly in favour of boys. This suggests that boys achieved at a higher level than girls. However, the effect size is trivial and t-values are non-significant. Thus, in all the comparisons the differences are neither practically nor statistically significant. It should be noted that Keeves (1968) using a very crude approximation for the design effect of the complex sample reported a significant difference. However, the

procedures employed currently in significance testing, make proper provision for the complex sample design using the jackknife routine in WesVarPC (Brick et al. 1997).

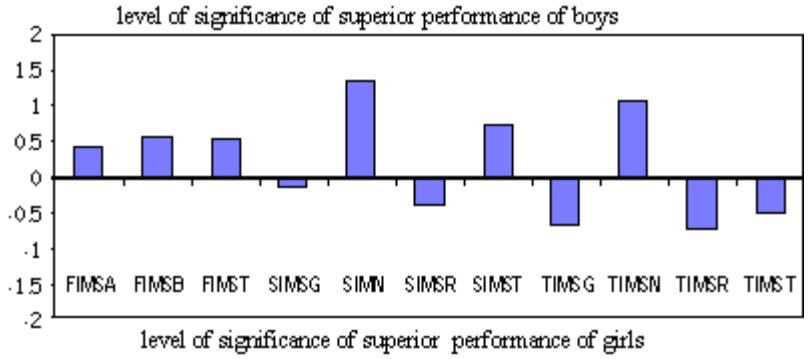


Figure 1. t-values for differences between boys and girls in Mathematics test score for FIMS, SIMS and TIMS

Comparisons of Sex Differences between 1978 Australian Students

In the SIMS data set four comparisons are undertaken and presented in Table 1 and Figures 1 and 2. The comparisons are between boys and girls in government schools, nongovernment schools, in all schools, and the restricted sample of schools.

The mean score differences between the two sexes in government schools (both Government and Restricted) indicated that girls achieved at a higher level than boys. However, the differences were neither practically nor statistically significant since the effect size and t-values were too small to be considered significant.

Furthermore, the estimated mean score difference between male and female nongovernment school students in SIMS is seemingly in favour of the boys, although the effect size (0.22) is small and the t-value (1.35) is non-significant. Thus, the difference is not statistically significant. Nevertheless, the achievement level of the boys is higher than that of the girls by approximately two-thirds of a year of mathematics learning in Australian schools (as estimated in the mid 1990s), although this result may be influenced by the particular schools selected in the analysis and the result is not significant statistically.

The last comparison in SIMS is between boys and girls in all schools. The mean score difference is apparently in favour of boys. However, the effect size (0.05) and t-value (0.72) are very small. Hence, the difference between boys and girls in 1978 is not practically or statistically significant across Australian schools. Out of the four comparisons, only in government schools is the achievement of the girls slightly higher than that of the boys. In the remaining comparisons boys appear to perform better than girls. These findings appear dissimilar to the findings in FIMS. In the 1964 data set all the differences are in favour of boys. This might indicate that over the 14-year period there was a small shift in achievement level differences between boys and girls. However, the differences are not statistically or practically significant.

Comparisons of Sex Differences between 1994 Australian Students

The estimated mean score differences between the two sexes in government schools (both Government and Restricted) are in favour of girls (see Table 1 and Figures 1 and 2). However, the effect size and t-values are very small and the mean differences are neither practically nor statistically significant.

The other comparison is between the two sexes in nongovernment schools. Unlike the government school students the difference is seemingly in favour of boys. However, the difference is not statistically or practically significant since the effect size is trivial (0.12) and the t-value is very small, because the design effects are large.

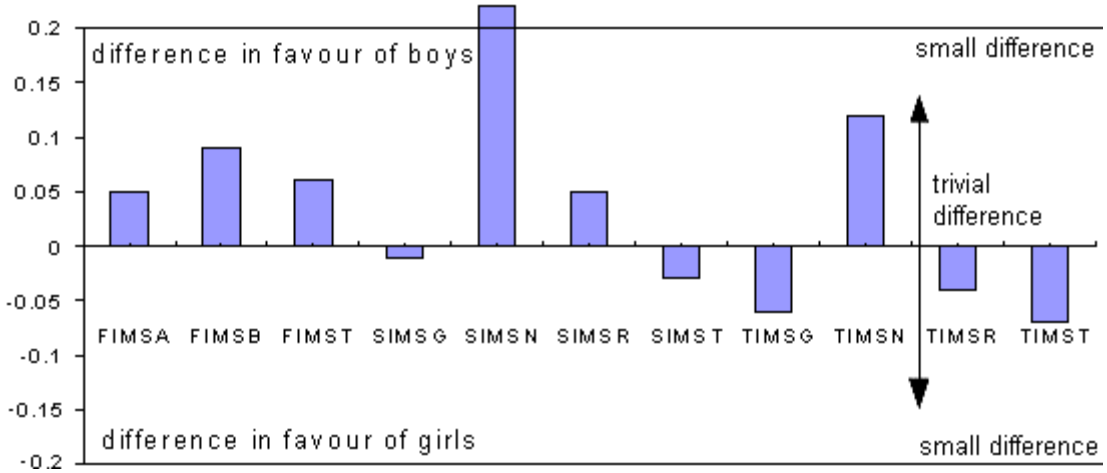


Figure 2. Effect size values for differences between boys and girls in Mathematics test score for FIMS, SIMS and TIMS

The last comparison in TIMS considers all students (government and nongovernment together). The mean score difference is apparently in favour of girls. However, the effect size (-0.04) and t-value (-0.52) are too small to be considered significant. Consequently the mean difference is neither practically nor statistically significant.

Three of the four different comparisons in TIMS between boys and girls in mathematics achievement reveal that the mathematics achievement levels of the girls are slightly greater than that of the boys. When the findings in TIMS are compared with the findings in FIMS and SIMS, the direction of the difference between the sexes has apparently changed, but the changes are not significant. In Figures 1 and 2 these results for the differences in the t-values employed for significance testing and in effect size are presented graphically in order to summarize the results.

In 1964 the differences are in favour of the boys, while in 1978 the difference is in favour of the girls only in government schools. However, in 1994 the differences, except in nongovernment schools, are in favour of the girls. Nevertheless none of the differences are found to be statistically significant. These nonsignificant differences might seem to suggest that some differences between the sexes in mathematics achievement are starting to emerge in the 1990s in favour of girls and in contrast to the findings of Keeves (1973), Carss (1980), Moss (1982), Leder and Forgasz (1992). These researchers would appear to contend that a difference between the sexes in mathematics achievement in Australian schools starts to emerge at the junior secondary school stage in favour of boys. The apparent change in direction in the 1990s could then be argued to have been a result of the implementation of different government policies to increase the participation and the mathematics achievement level of girls by the State and Federal Governments. Alternatively, it could be argued that the level of performance of the boys has declined more than that of the girls over time, as can be seen from Table 1 there is a general decline in achievement of all Australian students over the 30-year period. The possibility of a significant decline in the level of achievement in mathematics of boys, rather than a noticeable gain in the achievement of girls would be of some concern for Australian education.

Table 3. Descriptive statistics for mathematics achievement of male and female students for the three occasions

	FIMSA		FIMSB		SIMS		TIMS	
	Male	Female	Male	Female	Male	Female	Male	Female
Mean	462.0	457.0	455.0	448.0	440.0	443.0	422.0	431.0
Standard Deviation	101.0	90.0	87.0	76.0	105.0	99.0	126.0	126.0
Jackknife Standard Error	9.7	7.3	10.1	7.5	4.7	5.8	8.4	9.2

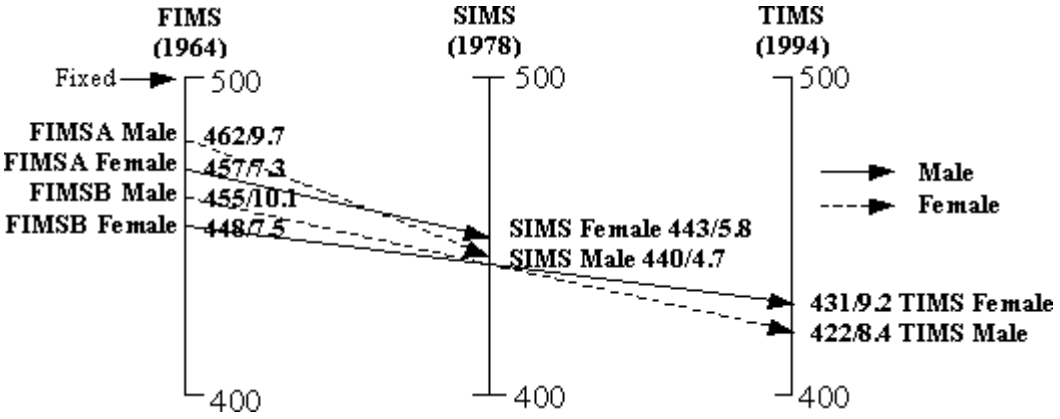
	Mean difference	Effect size	Effect as learning in	t-value	Significant level
FIMSA Male vs SIMS Male	22	0.21	5.5 months	2.04	<0.05
FIMSA Female vs SIMS Female	14	0.15	3.5 months	1.50	Not Significant
FIMSB Male vs TIMS Male	33	0.30	8.25 months	2.51	<0.05
FIMSB Female vs TIMS Female	17	0.16	4.25 months	1.42	Not Significant

SEX DIFFERENCES IN DECLINE OF ACHIEVEMENT OVER TIME

While this paper has so far reported no significant differences between the sexes on the three occasions, the possibility exists that there has been a significant decline in the standard of achievement of the boys over time, and no significant decline in the achievement of girls.

Table 3 and Figure 3 present the results of the analyses that test the differences in achievement in mathematics for boys and girls separately between occasions.

When the 1964 13-year-old male students (FIMSA) mean score is compared with the mean score of the 1978 13-year-old male students (SIMS), the mean score of the FIMSA students is higher than that of the SIMS students. The difference is 22 centilogits (see Table 3 and Figure 3). The difference is practically and statistically significant at 0.05 level. This significant difference shows that the mathematics achievement of male students declined over time by more than half a year of mathematics learning.



Scaled mean scores recorded with standard errors of mean.

Figure 3. The Mathematics test scale of government school students, FIMSA, FIMSB, SIMS and TIMS

The mean difference between FIMSA and SIMS female students is 14 centilogits (see Table 3 and Figure 3). This difference is not practically or statistically significant. This indicates that there is no significant difference between 1964 and 1978 female students mathematics achievement.

The mathematics achievement mean difference between 1964 Year 8 (FIMSB) and 1994 Year 8 (TIMS) male students is 33 centilogits (see Table 3and Figure 3). This difference is practically and statistically significant significant at the 0.05 level. This significant difference indicates that the mathematics achievement level of male Year 8 students declined from 1964 to 1994, by approximately one year of mathematics learning.

The mean difference between FIMSB and TIMS female students is 17 centilogits (see Table 3 and Figure 3). However, this difference is not practically or statistically significant. This indicates that there is no significant difference between 1964 and 1994 female students' mathematics achievement.

Hence, these four analyses show that the mathematics achievement of male students has declined significantly over the last 30 years, while there is not a significant decline for female students over time.

The reasons for the decline in achievement have not been examined. However, Tilahun and Keeves (1999) have shown that the Australian states differ in the extent of the decline recorded with one state out of the five showing a slight but not statistically significant gain in achievement. Nevertheless, it can be argued that any relative change in the performance in mathematics of girls with respect to boys is not a consequence of the improved performance of girls, but rather the significant drop in the performance of boys.

DISCUSSION AND CONCLUSION

In 1964 the differences are in favour of the boys, while in 1978 the difference is in favour of the girls only in government schools. However, in 1964 the differences, except in nongovernment schools are in favour of the girls. Nevertheless none of the differences are found to be statistically significant. These nonsignificant differences might seem to suggest that a difference between the sexes on mathematics achievement is starting to emerge in the 1990s in favour of girls and in contrast to the findings of previous studies. These researchers would appear to argue that a difference between the sexes in mathematics achievement in Australian schools starts to emerge at the junior secondary school stage in favour of boys. The apparent change in direction reported above might then be argued to have been a result of the implementation of government policies to increase the participation and the mathematics achievement level of girls by the State and Federal Governments. The evidence presented in this paper suggests that other forces are operating, that have resulted in a decline in performance of both male and female students of some practical significance when assessed in terms of a year of mathematics learning as well as being of a small but recognizable size for boys when assessed in terms of an effect size and statistical significance.

Until, there is clear evidence of statistically significant differences between the sexes in mathematics achievement and of sufficient magnitude to warrant resources being diverted to addressing the problem, there would appear to be more pressing problems to be considered in Australian education than the nature and origin of differences between the sexes in mathematics achievement at the lower secondary school level (see, for example, Rowe, 1998; Marsh and Rowe, 1996, particularly since both studies have ignored the multilevel nature of the data in their analysis). Moreover, such programs as have been undertaken in Victoria of establishing separate classes for boys and girls in mathematics at the lower secondary school stage would seem to be a misdirection of effort and resources until it can be shown that a substantial problem exists. Futher (1995) has already questioned the existence of such a problem.

However, this comment should not be taken to divert attention from the problem of the lower participation of girls in advanced mathematics classes at the senior secondary school level.

The failure to detect significant differences between the sexes in mathematics achievement in this study which uses the largest and most carefully designed samples available in Australia over a period of more than 30 years has several important implications for educational research in this field. There is a need to :

- employ large well designed random samples and to attain high response rates in order to test relationships that are widely assumed to exist;
- use appropriate statistical procedures to test for the statistical significance of differences that allow for the multilevel or nested nature of the data;
- report and discuss the magnitudes of effect sizes and the pattern of results, as well as use appropriate procedures to test for statistical significance;
- conduct research into the problems encountered by both boys and girls in learning mathematics as they pass through the lower and middle secondary stage;
- monitor in a systematic way changes in mathematics achievement over time, as well as changes in the performance of clearly identified subgroups of students; and
- maintain the view that all students can learn mathematics successfully irrespective of sex at all stages of schooling and in particular at the lower and middle secondary school levels.

ACKNOWLEDGMENT

The first author was sponsored by a Flinders University of South Australia Overseas Postgraduate Research Scholarship and the Flinders University Research Scholarship, while undertaking these analyses.

The authors would like to thank the several unknown reviewers of this paper for their helpful comments, which have been taken into consideration in the revisions of the paper.

REFERENCES

Adams, R. J. and Khoo, S.T. (1993). *Quest: The Interactive Test Analysis System.* Hawthorn, Victoria: ACER.

Baker, D. P. and Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education, 66*(2), 91-103.

Barnes, M. and Horne, M. (1996). Gender and mathematics. In P Sullivan, K. Owens and B. Atweh, (eds) *Research in Mathematics Education in Australasia 1992- 1995.* University of Western Sydney, Macarthur: Mathematics Education Research Group of Australasia. 51-87.

Brick, J. M., Broene, P., James, P. and Severynse, J. (1997). *A User's Guide to WesVarPC.* (Version 2.11). Boulevard, MD: Westat, Inc.

Byrne, B. M. (1989). *A Primer of LISREL Basic Applications and Programming for Confirmatory Factor Analytic Models*. New York: Springer-Verlag.

Bryk, A. S., Raudenbush, S. W. and Congdon, R. (1992). *Hierarchical Linear Models*. Newbury Park, CA: Sage.

Carss, M. C. (1980). Girls, mathematics and language: Some observations from classrooms. In Adelaide College of the Arts and Education (eds.).*Improving maths for girls: Report of a conference held at Raywood inservice centre, 30 May-1 June, 1980*. Adelaide:ACAE.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112 (1), 155-159.

Finn, J. D. (1997). Hypothesis testing. In J P Keeves (ed.), *Educational Research, Methodology, and Measurement: An International Handbook*, (2nd ed.), Oxford: Pergamon, 556-561.

Futher, R. M. (1995). The effects of single-sex classes on students' confidence and participation in Mathematics. Unpublished Med Thesis, University of Melbourne.

Hambleton, R. K.and Cook, L. L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement, 14 (2),* 75-96.

Jöreskog, K. L. and Sörbom, D. (1992). *LISREL8 and PRELIS2 for DOS, DOS Extender and Windows.* Chicago: Scientific Software Incorporated.

Keeves, J. P. (1968). V*ariation in mathematics education in Australia*: Hawthorn, Victoria: ACER.

Keeves, J. P. (1973). Differences between the sexes in mathematics and science courses. *International Review of Education, 19,* 47-63.

Keeves, J. P. (1992). *The IEA Study of Science III: Changes in Science Education and Achievement: 1970 to 1984,* Oxford, Pergamon.

Leder, G. C. (1992). Mathematics and gender: Changing perspectives. In D A Grouws (ed.). *Handbook of Research on Mathematics Teaching and Learning, New* York: Macmillan, 597-622.

Leder, G. and Forgasz, H. (1992). Gender: A critical variable in mathematics education. In B Atweh and J Watson (eds), *Research in Mathematics Education in Australasia 1988-1991,* Brisbane: MERGA, 67-95.

Lokan, J., Ford, P. and Greenwood, L. (1996). *Maths and Science on the Line: Australian Junior Secondary Students' Performance in the Third International Mathematics and Science Study.* Melbourne: ACER.

Marsh, H. W. and Rowe, K. J. (1996). The effects of single-sex and mixed-sex mathematics classes within a coeducational school. A reanalysis and comment. *Australian Journal of Education, 40(2),* 147-161.

Megarry, J. (1984). Sex, gender and education. In S Acker (ed.). *Women and Education: World Yearbook of Education,* London: Kogan.

Moss, J. D. (1982). *Towards equality: progress by girls in mathematics in Australian secondary schools.* Hawthorn, Victoria: ACER.

Rosier, M. J. (1980). *Changes in Secondary School Mathematics in Australia, 1964-1978.* Hawthorn, Victoria: ACER.

Rowe, K. J. (1988). Single sex and mixed-sex classes. The effects of class type on student achievement, confidence and participation in mathematics. *Australian Journal of Education, 32(2),* 180-202.

Sontag, L. M. (1984). Vertical equating methods: A comparative study of their efficacy. *DAI, 45-03B,* 1000.

Spearritt, D. (1997). Factor Analysis. In J P Keeves. (ed.), *Educational Research, Methodology, and Measurement: An International Handbook*, (2nd ed.), Oxford: Pergamon, 528-539.

Tilahun Mengesha Afrassa and Keeves, J.P. (1999). Changes in students; mathematics achievement in Australian lower secondary schools over time. *International Education Journal, 1(1),* 1-21.

Wright, B. D., and Stone, M. H. (1979). *Best Test Design: Rasch Measurement.* Chicago: Mesa Press.

Buildings and Betterment: Influences on the Design of State School Buildings 1900-1920

Mike Dillon
Department of Education, Training and Employment *dillon.mike@saugov.sa.gov.au*

Almost from the beginning of universal education, educationalists and reformers promoted a rationale that the school building and learning were linked. This rationale was widely accepted in the United States, England and in Australia and was significant in influencing the design of schools. This paper outlines the development of this rationale and then draws on South Australian School Inspectors' reports during the period 1900-1920 to show how this rationale was expressed in South Australian schools through a focus on school hygiene, appropriate aesthetic classroom displays and the development of school gardens.

school building design, school landscape and architecture, school inspectors

INTRODUCTION

The design of a new school is generally defined by its allocated budget, the number of planned enrolments, educational, and aesthetic needs. The design will need to address specific student and curriculum spatial entitlements and also legislative requirements (such as those required by building codes). The design of the school will need to be sufficiently flexible to meet both current and future curriculum initiatives, although often the future initiatives may be as yet, unknown. The design will be expected to support and enhance the educational priorities and goals for the school. Ideally, it is also be expected that the design will have an aesthetic component so that the school buildings are integrated and in harmony with the site. Bieler (1996, pix) has also noted;

> learning to appreciate beauty is also a part of the educational context. And the quality of the environment can have an effect on children…. as they develop in spaces built with harmony and human needs in mind. The beauty of spaces and forms and the attraction of the colours chosen, effective acoustics, carefully selected materials, well-designed lighting and green surroundings all make a contribution which few would dispute.

Few would argue with Bieler's articulation of the positive link between the quality of the school building and teaching/learning experiences. That is, the "better" the educational facilities, the more effective the learning experiences for students are likely to be. However, there has been little research undertaken on the relationship between school buildings and learning outcomes, still, the notion endures. However, the concept of linking educational outcomes with the school environment can be traced to the beginnings of universal education.

A significant development towards mass education had occurred in the late eighteenth century when Joseph Lancaster and the Reverend Andrew Bell both independently rediscovered the principle of monitorial instruction and successfully demonstrated that it could be applied to elementary education for the poor. Under the monitorial system, the number of potential students

could be very high and the cost of education per child significantly reduced over individual instruction (Burchell 1989, p.2, 5). As Katz (1976, p.393) has noted, there was practically no opposition to the introduction of mass education; in fact, education became compulsory only after attendance had become nearly universal. However, as public education became established and more widely spread, the problem of providing and designing educational buildings became a major issue for architects since both organisational and architectural solutions had to be found to address this new and growing issue. Profound changes in society brought about by industrialisation and mechanisation during the nineteenth Century made the necessity of solving the problem of the delivery of public education more and more apparent.

Roth (1966, p.24) noted there was practically no tradition of school architecture to begin with but that a rationale for the design of school buildings developed, as public education became more widespread. This rationale, which developed during the late nineteenth and early twentieth century, supported a connection between school buildings, student learning and the promotion of civic virtue in students and was significant in influencing school building design (Cutler 1989, Symes 1996). The arrival of mass education acted as a focus for interest in the school building as a site of possible social reform and, during the mid nineteenth Century, a number of architectural treatises were published by educational reformers such Barnard and Mann (in the United States) and by Robson (in the United Kingdom). These writings had common themes, proposing that school buildings should be conceptualised in a particular way; namely as "noble" structures that would support the inculcation of appropriate social values and transform children into virtuous literate citizens. For example, Barnard wrote that a school should be a temple "consecrated in prayer to the physical, intellectual and moral culture" which was committed to "leaving the strongest impressions of truth, justice and patriotism every child" (Uline 1997, p.195). Barnard also proposed that the school, in its arrangements and furnishings, was to be an "emblem of ethical and rational standards" (Symes 1996, p.88). In a similar vein, Mann wrote that schoolhouse design was closely connected to "the love of study…proficiency, health, anatomical formation and length of life. These are great interests…and suggest great duties" (Cutler 1989, p.4). Robson also believed the school building exerted an influence on a child's personal development. He wrote,

> … children… whose manners, morals, habits of order, cleanliness, and punctuality, temper, love of study and the school, cannot fail to be in no inconsiderable degree affected by the attractive or repulsive situation, appearance, outdoor convenience and in-door comfort, of the place where they are to spend a large part of the most impressionable part of their lives. (Robson 1874, p.6)

A major consequence of the work of the educational reformers was the professionalisation of public education and the invention of the belief that a special building (ie the school) was required to enable children to meet their moral obligations and reach their cognitive potential. Thus the school building was linked to educational theory and the curriculum, making them partners in the learning process (Cutler 1989, p.2). Symes (1996, p.88) reported that the school architecture dissertations were widely read in Australia and commented on the "striking continuity" between the writings of the reformers and those of educationalists, reformers and school inspectors in Australia at the time.

This exploratory study sought to investigate the extent of this "continuity" in regard to South Australian schools with a view to contributing to understandings concerning the growth and development of schools in South Australia in the early part of the twentieth Century.

RESEARCH FOCUS AND APPROACH

Very little material was found on the topic of the development of South Australian school buildings. There appears to be little data on the history of school buildings in South Australian and certainly nothing like Burchell's comprehensive account of the development of school architecture in Victoria. However, it was known that South Australian had a strong school inspectorial system at the time in which this exploration was focused (1900-1920). It was also known that School Inspectors' reports of their visits to schools were published in the South Australian *Education Gazette.* The *Education Gazette* was regarded as the key means of communication of Department policy to teachers in the State. These official publications provided not only administrative guidelines and policy, but also curriculum issues relating to teaching methods, programming and content of lessons. To paraphrase Rodwell (1992, p.107), in the *Education Gazette,* it is possible to catch insights of the "mood" of the Department and the various issues which were in the air at the time. The *Education Gazette* provided an account of what the Education Department wanted its teachers to do and to know. Inspectors' reports published in the *Education Gazette* therefore became an important means of explaining to teachers what their obligations were in delivering curriculum and managing the school environment. Accordingly, issues that the *Education Gazette* published between 1900 and 1920 were examined for references made by Inspectors to the link between school buildings and student learning.

SUMMARY OF RESULTS

This examination of the numbers of the *Education Gazette* published between 1900 and 1920 showed that in writing about school buildings, Inspectors considered two broad themes. The first related to the hygiene conditions within the school and the second to the aesthetic conditions of the classroom and the school grounds.

There are two key aspects associated with health and hygiene and its effect on schools and their organisation. The first concerned the creation of a healthy school environment to improve student learning. Teachers became responsible for the establishment and maintenance of a healthy school. This was done by training, via directives in the *Education Gazette* and supervision from Inspectors. Later, this became a broader issue in which health and hygiene in schools became linked with broader issues of community health. In South Australian this culminated in the introduction of a system of medical inspection in schools.

The initial focus regarding healthy school buildings related to the elimination of the school environment factors that impeded student learning. The need to improve lighting and the importance of ventilation in schools were frequently reported by Inspectors as significant in improving acquisition of knowledge. For example Inspector Burgan (*Education Gazette* 10 October 1906 p.228) wrote that an improvement in the lighting, heating and ventilation in the school and the provision of appropriate school furniture would improve the hygiene conditions "to produce the *mens sana in corpore sano,* ["healthy mind in a healthy body"] without which educational progress must be greatly retarded". This theme of hygiene was soon extended by other Inspectors to describe links between order and cleanliness and the improvement of teaching and learning. There are recurrent references in the *Education Gazette* by Inspectors to cleanliness, tidiness and neatness having a beneficial effect on students' minds and therefore being part of a sound teaching methodology.

Ultimately, these reported concerns were translated into changes in school design. For example, several South Australian schools built during 1908 were praised by Inspectors for the extent of natural lighting and ventilation. Also, later in 1908 a new design of school desk was introduced. Written and pictorial advice was given through the *Education Gazette* on the correct posture to be adopted by students in using this desk in order to alleviate health problems such as curvature of the spine, round shoulders, stooping, contracted chest and defective vision (*Education Gazette* 4 November 1908).

By 1910, the references to the hygiene aspects of school buildings had been incorporated into a broader context of community health and the "well-being of the Nation." (*Education Gazette* 14 June 1910, p.149) This meant there was an obligation for schools to adopt strategies that prevented the spread of disease into the community and which also assisted in promoting children's health. Children's health became the focus of the medical profession with schools becoming the focus for medical inspections. In fact, the *Education Gazette* of 7 October 1910 published an article that linked the three aspects of educational improvement, health and a system of medical inspection and advised that,

> the first essential step is the introduction of an effective system of medical inspection of schools in all the States … The work of the school doctor … should also include the effective and scientific management of school buildings and apparatus in respect to sanitation …

Such a scheme of medical inspection was adopted in South Australian in 1912 and was influential in affecting school design. For example, the report of the Medical Inspector in the 1916 South Australian Parliamentary Papers presented data in graph form showing the difference between visual defects in well and ill-lighted classrooms. In this report, the Medical Inspector also expressed concern with the need to improve ventilation and hygienic conditions in schools. The improvements in school buildings as a consequence of this initiative are well documented in subsequent Inspectors' reports and it is noteworthy that from 1921 onwards, Inspectors were formally required to report on lighting, ventilation, warming, sanitation and hygiene under separate headings.

In addition to the emphasis on hygiene, Inspectors also reported on and supported actions taken by teachers which made the classroom as attractive as possible. There are frequent assertions throughout the *Education Gazette*s 1900-1920 that the educative part of the aesthetic environment plays a significant role in child development. However, the decoration and adornment of the classroom had to be done in particular ways to create a favourable and moral impression upon pupils. Teaching aids and charts needed to be close at hand but not next to "good" pictures. Pictures needed to be of a particular kind; for example, Symes (1996, p.98) noted,

> …teachers were castigated …for use of unsuitable pictures of negative aesthetic value (like those supplied by the Northern Territory government or railway department) or for displaying unsightly pictures produced by local tradesmen.

Flowers were important in the classroom but needed to be arranged "artistically". As Inspector Whillas (*Education Gazette* 7 November 1906, p.248) recorded,

> …Pot plants in the windows and flowers on the mantel make the school beautiful, and children eagerly assist in this work. Good pictures are frequently found in the walls of the larger schools, but they should be kept scrupulously clean…The teachers are keenly alive as to the moral effect of clean and tidy rooms and surroundings. Many schoolrooms are very tastefully decorated with pictures and specimens of the children's handiwork.

That is, appropriate classroom arrangements addressed a moral imperative that was assisted by the provision of "tasteful decoration." This moral imperative also became a familiar theme in Inspectors' reports with appropriate classroom decoration actually described as "silent moral teaching." (South Australian Parliamentary Papers 1921, p.42)

Hyams (1988, p.120) has noted,

> the inspectorial system was the chief vehicle for achieving uniformity of classroom practice and regimentation of teachers in order to set and maintain educational standards…Inspectors noted provision of and adherence to rigid Education Department requirements.

As such, the recommendations regarding classroom decoration were dutifully carried out in schools across South Australian. Evidence of this influence and of the ways in which classrooms were organised and decorated can be seen in photographs (1) of classrooms of the time.

The "silent moral influence" was extended to include making the school grounds as attractive as possible, particularly the development of the school garden. The positive influence on children's progress was generated from the aesthetic effect of a good garden, from the labour expended by students in the development and care of gardens and also for the skills it developed which would be of future use to students. As Inspector Smythe (*Education Gazette* March 1901 p.50-51) reported,

> Besides the good moral effect and influence exercised on the minds of pupils by attention to these matters, the experience and skill acquired in the cultivation of plants and the treatment of soils must prove of considerable future benefit in after life.

The school garden emphasis was strongly supported by the Education Department and the *Education Gazette* included a regular column on gardening and school grounds. This proved extremely popular, as indicated by the increasing space devoted to issues associated with plants, trees and gardening advice in the Gazette. A clear demonstration of the importance of the school garden emphasis is found in the celebration of Arbour Day to encourage the value of cultivating school gardens. Again, the influence of School Inspectors in supporting the promotion of the school garden is evidenced in photographs of the time showing the Arbour Day ceremony (2).

The development of the school environment became a symbol of educational progress as Inspector Martin noted,

> The new buildings of today …must tend to the betterment of the child; but still a greater factor to the betterment of the child is the better school environments which have come and must continue to come. The decorating of the classrooms, the beautifying of the playgrounds and the making of garden plots are all educative. (South Australian Parliamentary Papers, 1915, Vol.3).

CONCLUSION AND FUTURE DIRECTIONS FOR STUDY

Cutler (1989) and Symes (1996) noted that by 1900 there were distinctive styles of architecture for public buildings, such as prisons, post offices, railway stations, hospitals, banks and asylums, and that these settings were both suitable for and symbolic of the purposes that were carried out within their walls. The design of these public buildings was intended to express the importance of the work that was carried out inside the building and also to increase its effectiveness. Schools were also participants in this trend and the development of the rationale that linked learning and the school building was a significant feature in influencing their design.

The advent of mass education generated the need for an "architecture" that had not previously existed. The school rapidly became an important community institution. Learning and the school building were seen to be linked and the school building itself was considered to play a positive part in the child's education. The reports of South Australian School Inspectors strongly echo this idea and support the concept of the school building representing a basis for the development of learning and moral development. This rationale was accepted and promoted through official documents in South Australian, with School Inspectors actively advocating for improvements in school buildings that would enhance the capacity for student learning. Initially this began as an emphasis on school building hygiene: ventilation, lighting and sanitation Later, the hygiene of the school building became part of a broader public health issue. These concerns manifested themselves in changes to school design; in the size and number of windows, the size of classrooms, and the provision of appropriate furniture. Concurrently the decoration of the classroom and the involvement of pupils in improving the school grounds were actively promoted as "silent moral influences" in students' educational development. In this way, the school building became an important feature in children's' lives, promoting hygiene, order, industriousness and moral welfare.

Regretably, the school building does not feature prominently in educational research. Yet as Churchill noted, "We shape our buildings, and they shape us." (Ellis 1988, p.19) There is significant scope for further historical research on school buildings, including in South Australian, the account of the development of South Australian school architecture. Any future studies in this field will assist in contributing to a better understanding of our society through its educational development.

The view that there was a link between the school building and student learning was assumed and accepted and few of today's teachers would argue with the notion. Nevertheless, this brief overview also highlights other important issues relating to school buildings: is there a definite (and positive) relationship between the school building and the acquisition of knowledge and skills? Do teachers and students work independently of the condition of school buildings? Do school buildings, designed in specific ways, enhance students' educational achievements? These are also matters for further research.

Notes

1. See for example, the following photographs in the Mortlock Collection: Pupils and Teacher, ca 1890 (B 46952); Classroom, ca 1910 (B 36574); "A" room, Norwood (B 17568); Classroom, Keilli (B 36210); Classroom, ca 1910 (B35520); Chemistry Lesson, ca 1913 (PRG 280/1/10/381).

2. See for example, the following photographs in the Mortlock Collection: Arbor Day, Houghton (B 31083); Arbor Day, Coomandook (B 59129); Arbor Day, Bakara (B 30026).

REFERENCES

Bieler, E., Gunvall, P., Hacker, M., Moonen, J.M. and Yelland, R. (1996) *Schools for Today and Tomorrow*. OECD France.

Brubacher, J.S. (1931, reissued 1965) *Henry Barnard on Education*. Russell and Russell, New York.

Burchell, L. (1980) *Victorian Schools: A Study in Colonial Government Architecture 1837-1900.* Melbourne University Press.

Cutler, W.W. (1989) Cathedral of Culture: The Schoolhouse in American Educational Thought and Practice Since 1820. *History of Education Quarterly,* 29(1) Spring.

Ellis, A. (1988) School Design Discussions. *Journal of Christian Education,* Papers 91 (April).

Hyams, B., Trethewey, L., Condon, B., Vick, M. and Grundy, D. (1988) *Learning and Other Things: Sources for a Social History of Education in South Australia.* South Australian Government Printer.

Katz, M.B. (1976) The Origins of Public Education: A Re Assessment. *History of Education Quarterly*, Winter, 381-407.

Robson, E.R. (1874) *School Architecture: Being Practical Remarks on the Planning, Designing, Building and Furnishing of School Houses.* London.

Rodwell, G.W. (1992) Historical Research in Education. In Cavanagh, D.M. and Rodwell, G.W. *Dialogues in Educational Research.* William Michael Press, Darwin.

Roth, A. (1966) *A New School Building.* Thames and Hudson, London.

Symes, C. (1996) Building the Queenslander: The Contribution of School Architecture to the Formation of the Child. *Queensland Review,* 3(2).

Uline, C.L. (1997) School Architecture as a Subject of Inquiry. *Journal of School Leadership,* 7(2) March.

Moral judgment of preschool children of the State of Kuwait

Fatima Nazar
Department of Foundations of Education, Kuwait University

This paper examines aspects of Piaget's ideas in a non-European culture and is unique in that no study has previously been attempted on preschool children from the State of Kuwait. Consequently, this study has been designed to determine the moral reasoning of preschool children from the State of Kuwait on the dimensions of moral realism and justice of Piaget's morality.

moral judgment, moral realism, justice, Piaget, preschool children

INTRODUCTION

Perhaps the most prominent contribution to our understanding of children's moral development has been made by Jean Piaget. Piaget (1932) described moral judgment as formative processes that follow on one another. He defined two ideal types: the heteronomous and the autonomous. The heteronomous type shows unilateral respect for authorities and the rules they prescribe. The heteronomous type of moral reasoning is characterised by moral realism. Moral realism is associated with 'objective responsibility' which tends to value the letter of the law above the purpose of the law. This is why young children are more concerned about the outcomes of actions rather than the intentions of the person doing the act. Thus, preschool children are inclined to rely on the magnitude of the material damage caused by the actor, and not on their intrinsic motivation (Subbotsky 1993, p.77). In the same way it is difficult for the preschooler to understand the concept of an accidental wrong. It may be impossible to convince the young child that a classmate accidentally knocked down his block building or stepped on his toe because the area was crowded (Vasta, Haith and Miller 1995).

The autonomous type of equality and respect established at 10 to 12 years of age, is characterised by the ability to consider rules critically and selectively apply these rules based on a goal of mutual respect and cooperation. Human actions are judged by intentions and motives as well as consequences of actions. Children no longer rely on the magnitude of the material damage, but give priority to the story characters' intentions. They are now motivated in their relationships to others by mutual respect for others. Piaget (1932) stressed the benefit for children having experiences of interacting with others on an equal footing. He thought that a prime vehicle for such experience was discussion among peers.

> As the child grows up the prestige of other children diminishes … he can discuss matters more and more as an equal and has increasing opportunity of freely contrasting his point of view with that of others. (Piaget 1965, pp.95-96)

From Piaget's point of view, the child 's concept of justice thus, centres on a social system where there is balance and coordination of the interests of the individuals who are participating in it. In other words, each is given his fair share.

Since Piaget‘s seminal work, much research has been conducted from the Piagetian perspective on children from Western-countries to examine their moral judgments of damage done under accidental and intentional conditions (Johnson 1962; Ozbek and Forehand 1973; Flynn 1984; Berg-Cross 1971; Gutkins 1972; Keiten and Garg 1993). To date, however, there is a striking paucity of research on preschool children from Arab cultures. To this author’s knowledge, no study has been attempted on preschool children from the State of Kuwait. Consequently, this study has been designed to determine the moral reasoning of preschool children from the State of Kuwait on the dimensions of moral realism and justice of Piaget’s morality. Specifically, the following research questions are addressed in this study.

1. To what extent preschool children at the ages of five and six use intention and outcome information in judging other’s deeds?
2. Are there differences between boys and girls in both judgments in terms of motive and magnitude of physical damage (intrinsic-externic dimension)?
3. To what extent preschool children use equality or authority/obedience justification in their judgment about justice?
4. Are there gender differences in making judgments about justice issues (justice-dimension)?

METHOD

Subjects

Subjects were 108 preschoolers from private kindergartens. They were equally divided into sex and age: 27 boys and 27girls aged 5 years; and 27 boys and 27 girls aged 6 years (mean age = 5.3). All the children were Muslims from the State of Kuwait. They were contacted through their schools. Information about age and sex were taken from the school record.

Material

Four moral items were chosen for this study. They consisted of a pair of short stories patterned after Piaget’s clinical method (1932) to assess the children’s moral realism; and two moral items dealing with justice. The four stories were derived from Piaget’s work which intended to be a representative subset of the stories that Piaget reported in his work on moral realism (Piaget 1932, p.117) and on the development of the idea of justice (Piaget 1932, p.195). The moral realism stories deal with the child’s ability to consider subjective intention when making moral judgments about clumsiness/stealing. The justice stories contained questions involving equality versus authority. The four stories and probe questions were translated by the author into Arabic without changing any of the content except of the Christian names were replaced by Muslim names; John and Henry were replaced by Ahamd and Khalid (masculine version) and by Aisha and Fatima (feminine version):

The following are the pair of stories dealing with moral clumsiness/stealing:

1. John was playing in his room when his mother asked him to come to dinner. While John was walking by the table, he accidentally slipped and bumped the dishes: 15 dishes fell and broke.
2. One day, when Henry’s mother was not home, he decided to eat some cookies even though his mother told him not to. While he was opening the cupboard to take some cookies, one dish fell and broke.

Which of these boys did the worse thing? Why?

The following two stories deal with authority/obedience (justice):

3. One afternoon, a mother asked her two daughters to help around the house because she was tired. She asked one girl to dry the plates, and the other to set the table. One of the girls decided she did not want to help and went outside to play. The mother asked the other girl to do both of the chores.

What did she do?

Was it fair of the mother to ask her to do both chores?

Why or why not?

4. One day a father asked his two sons to wash the car because he was tired and could not wash it himself. He asked one boy to wash the outside of the car, and the other boy to clean the inside of the car. One of the boys did not want to clean the car, and went to play with his friends. The father asked the other boy to do both the washing and the cleaning of the car.

What did he do?

Was it fair of the father to ask him to do both the washing and cleaning?

Why or why not?

Coloured drawings were used to draw the children's attention to the presented stories and to render the task more understandable and more concrete. It may be useful to note here that the cultural items of the stories seem suitable for the children's culture, and therefore, they translate well into that cultural context.

Procedure

After a brief contact with each child to establish rapport, the author, a trained examiner with the preschool children, and whose language was that of the child, interviewed the children individually. Prior to assessing the subjects' moral judgments, they were asked to recount the pair of stories in their own words. They did not have any difficulty in grasping the stories. The answers were transcribed verbatim for further analyses. The scoring of the responses to the mortal stories was done on the basis of a scoring system similar to the one presented by Lourenco (1991). Responses were classified according to the outcome of an act, or the intention of an act (stories 1 and 2); and according to the justification for the fairness or unfairness of an act (stories 3 and 4). Percentages of answers for each category were calculated on the basis of their frequencies.

RESULTS

Children's evaluations of the moral items related to moral realism

The results of the study indicate that 74 children (69%) regarded the protagonist who broke one dish as having done the worse thing, whereas 34 children (31%) opted for the protagonist who broke 15 dishes as having done the worse as displayed in the Table 1.

Thus the sample as a whole prefers the significant use of the intention of the actor rather than the magnitude of damage (Chi-square = 64.8, $p < 0.05$). The older children (6 years old) made significantly more moral judgments in terms of intrinsic motivation (80%) and less moral judgments in terms of physical damage (20%). Whereas, younger children (5 years old) did the opposite; (57%) made less judgments in terms of intrinsic motivation, and (43%) made more

judgments in terms of physical damage (Chi-square = 24.9, $p < 0.05$, and Chi-square = 41.3, $p < 0.05$) respectively as displayed in Table 2.

Table 1. Frequency and percentages of responses in terms of physical damage and intrinsic motivation of the whole sample (n = 108)

Responses	n	%
Physical damage	34	31
Intrinsic motivation	74	69

Table 2. Frequency and percentages of types of judgments according to age

Responses	5 years	6 years
Physical damage	23 (43%)	11 (20%)
Intrinsic motivation	31 (57%)	43 (80%)

In order to determine whether sex of children is related to their judgments a series of chi-squares was performed. For the younger children (5 years), females made significant moral judgments in terms of intrinsic motivation (24%) than boys (15%). Chi-square =20.6, $p < 0.05$. Similarly, for the older children (6 years), females made more significant moral judgments in terms of intrinsic motivation (44%) than boys (26%) (Chi-square = 46.3, $p < 0.05$). However, concerning judgments in terms of physical damage, the reverse was true. For the younger children (5 years), boys made more significant moral judgments in terms of physical damage (48%) than females (30%) (Chi-square = 41.3, $p < 0.05$). In a similar vein in the older children (6 years) group, males made more significant damage (22%) than females (7%) (Chi-square = 71.0, $p < 0.05$).

Children's evaluations of moral items related to justice

The results indicate that the majority of the children said it was not fair that the protagonist should do both of the chores (89%), because the second protagonist should also do his part (Equality orientation). However, only 12 Children (11%) responded that it was fair because the mother/father said so (Authority/obedience orientation). Table 3 displays the results of the children's evaluations and justifications of their judgments according to age.

Table 3. Children's evaluations and justifications related to moral items about justice according to age

Type of judgment and justification	5 years	6 years
Not fair (Sister/Brother should help) (Equality Orientation)	45 (83%)	51 (94%)
Fair (Mother/Father said so) (Authority /obedience)	9 (17%)	3 (6%)

The Chi-square test performed on these findings indicate no significant differences between the two age-groups in terms of their judgments related to justice (Chi-square 1.7, $p < 0.05$). However, concerning their judgments of justice in terms of obedience/authority, the Chi-square test revealed a significant difference between the two age-groups (Chi-square: 85.7, $p < 0.05$). It should be noted here that the children who made judgments related to justice in terms of authority/obedience were all females (n = 12).

DISCUSSION

Children's evaluations of moral items related to moral realism

The results of this present study reveal that the subjects' moral judgments were made predominantly in terms of intrinsic motives (69%). However, their orientation towards moral judgments in terms of magnitude of physical damage (31%) should not be ignored. These findings are consistent with Flynn's (1984) study which indicate that preschool children are capable of making moral judgments in both apology-restitution and guilt-innocence. Likewise, in this study preschool children made judgments in terms of both intrinsic motivation and amount of physical damage.

Results by chronological age are quite consistent with Piaget's theory which suggests that immature moral judgments reflect centring on consequences while disregarding intent. In contrast, mature responses reflect a recognition that both intent and consequence are important in solving moral dilemmas (Foye and Simeonsson 1979). Thus, older children (6 years) made more significant judgments in terms of intrinsic motivation than younger children (5 years). On the other hand, younger children made more significant judgments in terms of magnitude of physical damage. This finding corroborates research that suggests that chronological age substantially influences moral development in children (Johnson 1962; Ozbek and Forhand 1973).

The findings that boys develop moral judgments earlier than girls was not consistent with Piaget's (1932) suggestion that boys develop moral judgments earlier than girls because of differential demands of peer cooperation. Gender differences in moral development has always been an issue of debate in the research literature. A lot of inconsistent statements have been made as to the effect of the gender on moral development (Huston 1983; Lytton and Romey 1991). Consequently, a clear verdict cannot be given in favour or against a gender difference issue, since consistency has not been established yet. The findings of this study seem to support the assumption that children around the age of five years judge an actor who had no bad motives, not to be bad, although he had unintentionally caused damage. They base their judgments on the criterion of a match between an actor's motive and the outcome of his action (Nunner-Winkler and Sodian 1988). Piaget contends that moral judgment begins to be made at the age of six or seven after a year of socialisation in school. Children in this study, however, attended kindergarten schools, which may have enhanced the socialisation process and hence the early emergence of their moral judgment. A common feature in these schools concerns the social atmosphere that seems conducive to children's moral growth. Curriculum activities provide opportunities for children to judge moral events they encounter in their interactions with their peers. Children are exposed to many activities that involve them in discussing stories and puppet skits presented to them from time to time as part of the curriculum. It is quite interesting to mention that children in these schools are introduced early to explore the concept of intention and motive using stories and puppet skits to discuss the characters' motivation. Dramatic play and role play are also encouraged. Such activities surely enable the children to stand on another person's shoes and promote their development of perspective taking. Hence, it may not be surprising that children display some surprisingly sophisticated moral judgments at this age.

Children's evaluations of moral items in terms of justice

Concerning the children's judgments on the issue of justice, the results of the present study indicate that the majority of the children 96 (out of 108) clearly indicated the idea that it was not fair of the mother or father to treat the children unequally. Both children should do their part of the chores. They all responded in terms of equality. These data indicate that the children begin to

give up their moral realism and to make judgments consistent with the notion of justice, equality well before the age of 7. Piaget reported, by means of stories similar to the ones used in this study, that about 75 per cent of the children of 5 to 7 years defend obedience; and that about 80 per cent of the children between 8 and 12 years defend equality (Piaget 1932, p.268). The present finding is not consistent with Piaget's original findings. This may be due to the early socialisation process of the preschool children in the sample. Only nine children responded in terms of obedience. They identified what is just with what is in conformity with obedience. This finding illustrates well the differences of children's judgments as a function of sex. All the girls reported that the mother was fair because she said so (authority/obedience). This result may be interpreted in terms of socialisation processes in Muslim families where it is stressed that girls should be more obedient than boys. This finding also illustrates that the issue of gender differences in moral judgment might be the result of many factors. It is differentially influenced by social-cultural environment, and hence the inconsistent findings reported in the literature concerning this issue (Smetana 1981, 1985; Smetana and Braeges 1990).

In conclusion it might be interesting to attempt to link the main findings of this present study to current research and debate in this area. A major finding of this study is that Piaget's time-line for moral development of Muslim Kuwaiti children apparently does not hold. Instead, children in this study had greater moral reasoning ability at a younger age than children in Piaget's study. This result is consistent with findings of the proponents of domain theory. Domain theorists contend that young children's thinking in the moral domain is actually quite sophisticated (Smetana 1995; Tisak 1995) and some even claimed that early moral thoughts should be characterised as "moral, not premoral" (Killen 1991, p.115). Domain researchers have shown that "young children do make moral judgments that go beyond heteronomous obedience to authority and rules" (Turiel 1983, p.148). They have equally shown that by the age of five years children have the competence to think of moral issues as "obligatory, generalisable, and impersonal" (Turiel, Killen, and Helwig 1987, p.169). In this vein, there may be no surprise that in this present study five and six year old children had greater moral reasoning ability at this early age. Thus, the major finding in this study seems to support the claim of domain theorists and challenge Piaget's account of early morality.

REFERENCES

Berg-Cross, L.G. (1971) Intentionality, degree of damage, and moral Judgments. *Child Development*, 46, 970-974.

Flynn, T.M. (1984) Age, sex, intelligence, and parents' occupation and the moral development of the preschool child. *The Journal of Psychology*, 117, 251-255.

Foye, H. and Simeonsson, R.(1979) Quantitative and Qualitative analyses of moral reasoning in child, adolescents and adults of similar mental age. *Journal of Pediatric Psychology*, 4, 197-209.

Gutkins, D. (1972) The effect of systematic story changes on intentionality in children's moral judgments. *Child Development*, 43, 187-195.

Hebble, P.W. (1971) The development of elementary school children's judgments of intent. *Child Development*, 43, 187-195.

Huston, A. (1983) Sex typing. In E.M. Hetherington (Ed.), *Mussen Handbook of Child Psychology* (Vol.14), New York: Wiley.

Johnson, R. (1962) A study of children's moral judgments. *Child Development*, 33, 327-354.

Keilen, M.V. and Garg, R. (1993) Moral Realism in adults' judgments of responsibility. *The Journal of Psychology*, 12(2), 149-156.

Killen, M. (1991) Social and moral development in early childhood. In W.M. Kurtines and E. Gewirtz (Eds.) *Handbook of Moral Behaviour and Development: Research* (Vol.2), Hillsdale, NJ: Erlbaum.

Lourenco, O.M. (1991) Children's distinction between beauty and morarlity. *Archives de Psychologie*, 59, 247-261.

Lytton, H. and Romnay, D.M. (1991) Parents' differential socialisation of boys and girls: A meta-analysis. *Psychological Bulletin*, 109,267-297.

Nunner-Winkler, G. and Sodian, B. (1988) Children's understanding of moral emotions. *Child Development*, 59, 1323 -1338.

Ozbek, N. and Forehand, R. (1973) Factors influencing the moral judgments of retardates. *Journal of Mental Deficiency Research*, 17, 255-261.

Piaget, J. (1932) *The Moral Judgment of the Child*. Routledge and Kegan Paul.

Smetana, J.G. (1995) Morality in context: Abstractions, ambiguities and applications: In R.Vasta (Ed.), *Annals of Child Development* (Vol.10), London: Kingsley.

Smetana, J.G. (1985) Preschool children's conceptions of transgressions: The effects of varying moral and conventional domain-related attributes. Developmental Psychology, 21, 18-29.

Smetana, J.G. (1981) Preschool children's conceptions of moral and social rules. Child Development, 52, 1333 - 1336.

Smetana, J.G. and Braeges, J. (1990) The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly*, 36, 329 -346.

Subbostsky, E.V. (1993) *The Birth of Personality: The Development of Independent and Moral Behaviour in Preschool Children*. Harnester.

Tisak, M.S. (1995) Domains of reasoning and beyond. In R.Vasta (Ed.) *Annals of Child Development* (Vol.11), London: Kingsley.

Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. New York: Cambridge University Press.

Turiel, E., Killen, M. and Helwig, C.C. (1987) Morality: Its structure, functions, and vagaries. In J. Kagan and D. Lamb (eds.) *The Emergence of Morality in Young Children*, Chicago: University of Chicago Press.

Vasta, R., Haith, M. and Miller, S. (1995) *Child Psychology: A Modern Science*. New York: Wiley.

Zahn-Waxler, C., Radke-Yarrow, M. and Brady-Smith, F. (1977) Perspective-taking and prosocial behaviour. *Developmental Psychology*, 13, 87-88.