Proceedings of the Educational Research Conference 2002
Special Issue: Volume 3 Number 4
November 2002

Published by

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

International Education Journal

Volume 3, Number 4, November 2002

Educational Research Conference 2002 Special Issue

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 blind peer-review process by two experts in the field.

Published by:
Shannon Research Press,
39 Dorrien Avenue,
Woodcroft, SA 5162,
Australia

ISSN 1443-1475

Designed by Katherine Dix

Printed in Adelaide
November 2002

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.

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 dissertations as occasional monographs. 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.

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 that 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.

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

Submissions should be forwarded electronically, either by:

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

or,

on a 1.4MB 3.5 inch floppy disk, including a printed copy, to:

IEJ Online Editor,
39 Dorrien Avenue,
Woodcroft, SA 5162

Manuscript Details

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 exceed 150 words
- paper should not exceed 5,000 words (20,000 for occasional papers)
- include 5 key words describing the article including those in the title
- be in final form ready for reviewing
- conform to the APA Reference System

General Principles for layout:

- Papers should be prepared using the downloadable template *iejbase.doc* with Microsoft WORD on Mac or PC.
- Use the heading and formatting styles contained in the template, under the Styles Menu, to format article.
- Images and tables should be correctly positioned with caption so that they are referred in the text before their appearance.
- Include images that are imported or inserted into the text from an external source as separate graphic files (GIF or JPG)
- Figures constructed in WORD from the Draw Menu must be 'grouped'.
- Use footnotes, 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 *iejbase.doc*, just modify appropriately.
- Do not use page numbers in text.

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.

FUIIE web site:
www.flinders.edu.au/education/fuiie

Contemporary Themes in the Research Enterprise

Tony Gibbons
Flinders University *tony.gibbons@flinders.edu.au*

Gavin Sanderson
Flinders University *gavin.sanderson@flinders.edu.au*

This paper discusses themes associated with the enterprise of educational research, particularly as it applies to notions of ontology and epistemology in 'quantitative' and 'qualitative' methodologies. It begins by identifying the 'space' that educational research occupies in the contemporary social setting and indicates the growing complexities of the 'real world' which researchers contemplate. The development of empirical and interpretive methodologies is traced through an historical analysis of educational research which highlights the emergence of the postmodern platform. This leads to an examination of the traditions of both the 'quantitative' and 'qualitative' approaches with a view to understanding their positioning in the field of educational research. It is maintained that these two methodologies are examples of paradigms which exist within traditions. They are often said to be incommensurable paradigms. We propose a solution to their seeming incommensurability which in part involves assessing the coherence of the idea of a 'paradigm'.

Research, Qualitative, Quantitative, Paradigm, Tradition

INTRODUCTION

> Oh East is East, and West is West, and never the Twain shall meet
> Till Earth and Sky stand presently at God's great judgement seat
> But there is neither East nor West, Border no Breed, nor Birth
> When two strong men stand face to face, though they come from the Ends of the Earth[1]

The opening stanza of Kipling's work 'The Ballad of East and West' provides a paradox that sets the scene for this paper. Initially, the poem describes two camps that are seemingly eternally separate and anathema to one another. In the absence of any commonality, a schism separates them. Later, this 'difference' is cast aside and is replaced by a perception of 'sameness' when social and historical identifiers are stripped away by the process of peoples from the East and West coming to face each other. Their shared humanity unites them (despite the inference of a lingering propensity for hostility). We are left feeling that East and West can be simultaneously different *and* the same, depending upon the perspective of the observer. So, too, can this analogy be extended to the enterprise of educational research.

[1] Poet's Corner (2001) Rudyard Kipling, *The Ballad of East and West*. Available on-line: http://www.geocities.com/~spanoudi/poems/kiplin01.html [2002, 28 August]. The poem is used as a vehicle to introduce the theme of the paper. 'East-West' should not be interpreted as suggesting 'global coverage' in the sense that there are no methodologies other than those which are either 'qualitative' or 'quantitative'.

On one hand, there is talk of a great divide separating 'quantitative' and 'qualitative' research, based on the existence of two perspectives on the study of human behaviour. This has profound implications for how educational research is conducted (Burns 1990, p. 1). It has been argued that proponents of the 'P-Theory', or the idea that we are dealing with two different research paradigms come in two classes. There are those who subscribe to the *'oppositional diversity thesis'* and maintain that the two paradigms are epistemologically different and incommensurable. Others are sympathetic to a *'complementary diversity thesis',* which suggests that whilst they may be epistemologically distinct (and incommensurable), they overlap in their different approaches to enquiry and are complementary, not competitive. A position against the P-Theory, the *'unity thesis'*, denies that there is fundamental diversity in educational research and that a commonality of purpose means that there is no sound epistemological basis for having separate research 'paradigms' (Walker and Evers 1997, p.22). This paper is critical of the idea that paradigms are incommensurable and takes issue with the notion that quantitative and qualitative research are best described as paradigms. In our view they are traditions, which contain a number of paradigms. In fact Walker and Evers talk of them as traditions and confuse the issue of the distinction between traditions and paradigms. It is our view that the two traditions, while distinct, can and should be brought together in educational research (Keeves 1997, p.7)

For all researchers, it is important to have some understanding of the issues raised here, if only to better understand how their research activities may be justified. At the moment there is confusion in educational research engendered by the different approaches. Staff and students alike have been known to hoist a methodological flag over their research territory and defend it with vigour against claims that their particular way of conducting investigations is inappropriate, lacking in rigour, and 'not proper' research. This paper argues that most of the hubbub is less to do with enlightened dialectic, and more a case of a lack of familiarity with, and understanding of, the philosophical issues underwriting the divisions so often encountered between the so-called 'research paradigms'. Discussions of differences, if it is to be valid and constructive, must appeal to ontological and epistemological considerations; not a dismissive attack on those who support or eschew 'quantitative' over 'qualitative' and vice versa. The discussion in this paper goes beyond the theses expounded by Walker and Evers and suggests that although quantitative and qualitative approaches have their own strengths and limitations, they can be said to be as incoherent as each other when it comes to helping form universal generalisations about the 'real world'.

THE CONTEMPORARY SETTING

Initially, it is crucial to recognise the complicated milieu in which contemporary educational research takes place, for it would appear that the complexities associated with 'what and who are researched' have accelerated in an exponential-like manner in the past few decades. As civilisations across the globe enter the third millennium, never before has the planet been host to such a large and increasingly mobile human population. The population is presently climbing through six billion, having doubled in the last fifty years, and it is expected to pass nine billion by the middle of this century (US Census Bureau 2002). Never before has technology been so advanced and intertwined into the daily lives of so many people. The amount of information that most individuals can access and assimilate on a daily basis is unprecedented. Never before has human activity been able to so markedly leave an imprint made distinctive by its ubiquity and the way in which it both dominates and degrades the physical environment which sustains us. There are real questions about whether we can

survive as a species, given the damage that we are causing the Earth through overpopulation and a rapacious *growth-as-development* mentality. Never before has humanity been so materially rich, yet so poor, with a minority of the world's population controlling the majority of its wealth. Whilst some may argue that global inequality in personal incomes has fallen in real terms since 1975 (Bradford-DeLong 2001), it is clear that the domination of the world's resources by the developed counties continues to perpetuate major inequalities. In summary, "after a century of the greatest flowering of human knowledge, there are more poor, more knowledge-deprived, more suffering, more unsustainable development, more sick and dying than ever before" (Cribb 2002, p. 29).

To paraphrase Charles Dickens, today we are simultaneously witnessing the best of times and the worst of times! This is the landscape, which the enterprise of educational research presently contemplates, and, to extend the challenge, the changes in the next one hundred years are likely to exceed those of the last one thousand years in terms of impact, speed, scope and importance (Beare & Slaughter 1995, p. 5). This is likely to have significant import not only for the foci of research activities in the immediate future, but also for the *activity of research* in the 21st century, itself a contested terrain underwritten by a milieu which has inherited an economic, political and epistemological 'zeitgeist' from the last decade which was "considerably less certain and confident than it had been some thirty years previously" (Welch 1999, p. 35).

THE RESEARCH ENTERPRISE IN THE CONTEMPORARY SETTING

By the late 1990s, a plethora of approaches characterised the contemporary research enterprise (see Table 1). Keeves (1997) acknowledges the diversity by stating that;

> "…there is now a greatly increased variety in the strategies and tactics employed in research into educational problems, as well as in the methods, theoretical perspectives and analytical procedures that are being used to investigate the processes and practices, the context and conditions, and the products and policies which occur in the field of education" (p. xv).

Carspecken (1996) concurs by musing that a room filled with social researchers would be a cacophony of cliques, with each exhorting their own distinctive jargon and cultural style (p. 1). Such is the growing diversity and complexity in approaches; an indication that "the frontiers of educational research are constantly changing" (de Landsheere 1997, p. 15).

The enterprise of current educational research includes the largely scientific, or quantitative, approach[2] that is derived from natural science and was *de rigueur* throughout most of the 1900s, as well as the newer perspectives and methods offered by humanistic, or qualitative[3], researchers since the early 1970s. Both methodologies seek to make contributions to the 'body of knowledge' that allows us to use generalisations to benefit educational and social practices (Keeves 1997, p. 3; Arnove & Torres 1999, pp. 4-6). In addition to these two well-known positions, others have emerged which examine the workings of our societies, particularly in terms of relations of power and its consequences and clearly have their origins in Nietzsche. These are 'critical theory' or 'critical action research' which is directed at social change (Keeves 1997, p. 6), and the 'postmodern' approach, sometimes called 'poststructural' or 'deconstructive', which is a method of discourse analysis that analyses knowledge in terms of "who speaks, for whom and by what authority" (Smith 2000, p. 10). Both of these latter approaches are aimed at the politics of emancipation and social justice and resonate more with qualitative approaches than quantitative ones.

---

[2] Also referred to as scientific, positivist, empirical, logical empiricism, or objectivist approaches.
[3] Also referred to as interpretive or subjectivist approaches.

Table 1: Approaches to Enquiry[4]

Quantitative	Qualitative	Criticalist	Deconstructive / Poststructuralist
Approach			
Classical physical sciences investigation	Historical & existential studies valuing subjective understanding of subjects	Marxist, interpretive & psychoanalytic studies which focus on insights & judgements of subjects	Anthropological, psychoanalytic & linguistic understanding of the interrelationships between culture, language, desire & the self. Understands the self/subjectivity as decentred and in process
Assumptions about reality			
Reality is unitary & only understood by empirical analytic enquiry	Multiple realities exist & require multiple methods to under stand them	Multiple realities made problematic by distorted communication	No reality or 'real world' accessible beyond language. Reality constituted in & through discourse. Constituted by language & naturalised through ideology
Foundation of data			
Disciplined, rule-governed sensory-perceptual observation, i.e. rules for observation	Meaning is the basis for data & precedes logic & fact	Meanings found in language & social behaviour & precedes logic & fact	In language understood as discourse or a system of meaning. No fixed meaning because it shifts according to context & motivations of speakers & writers and listeners & readers
Observation			
Via clear & unambiguous rules, not modified by the setting and independent from it	Through social, linguistic & cognitive skills of researchers, i.e. dialogue	Interpretive methods & critical self-reflection concerning grounds of observation	Interrogating various discourses which constitute the field of enquiry & analysing power relationships generated through discourses
Outcomes of enquiry			
Evidence and generalisable laws not affected by context or the investigative methods. Objectivity removes error & bias	Knowledge that is dependent on the process of discovery. Integrity of findings based on quality of social, linguistic & cognitive skills of the researcher in data collection, analyses & conclusions	Knowledge which falls within the interpretive framework but also assists personal liberation & understanding & emancipation from forces constraining rational independence of individuals	How knowledge is constructed. Questions foundations & frame-work of knowledge. Asks how knowledge has been constructed as truth and how social realities are constructed through language
Inherent interests			
Prediction & control, technically exploitable knowledge. Explanation	Discovering meanings and beliefs underlying actions of others. Understanding at the level of ordinary language & action	Interpretive interests plus revealing interests underlying other forms of enquiry & action. Improving human existence. Practical outcomes for the public good	Questions totalising or unified interpretations & understandings. Views them as partial. Seeks to locate dominant interest & modes of producing & maintaining them. What positions are possible for marginalised groups
Inherent values			
Science & scientific knowledge are inherently value- neutral	Science & scientific knowledge must be interpreted in terms of the values they represent	Science & knowledge are never value-neutral. They always represent certain interests	Scientific truths & knowledge are never value-neutral; they are the effects of power. 'Value' is a contested terrain because of (i) inherent contradictions between the researcher & the researched, and (ii) contrary values within oneself due to ideologies of gender, race, age, ethnicity, class sexual preference, nationality, etc

Initially then, the challenge for the researcher would appear to be where to locate themselves in terms of a specific approach to their research and, more often than not, this is probably done by what they know in terms of their training as well as following the lead of their associates, peers, colleagues, and 'experts' in the field (Paul & Marfo 2001, p. 527). Siding with a particular methodology because it is favoured by a certain group of scholars, a funding body, the government, or wider society is understandable if not always justifiable.

[4] Adapted from Connolle (2000), pp. 32-33.

The starting point for any researcher should be for them to take time to reflect on the world that they know and ask "*of the things that I believe, why do I see them as such and what is the philosophical framework that makes it so?*". Honesty is not simply a social virtue but an intellectual virtue that demands that the researcher inspect their personal ontological and epistemological framework and see it in the context of their history in the society of which they are a part. This society is not simply the broad society in which they live but the community of researchers of which they are a part. This task is not easy and participation requires a judicious mix of openness, detachment, honesty, and logic, underwritten by a self-referential consistency and a willingness to step outside what is acceptable, even fashionable, and perhaps expected. In other words, the intellectual virtues. Therein lies the basic challenge and one that individuals and educational institutions must address in terms of "the professional preparation of educational researchers" (Paul & Marfo 2001, p. 534).

BACKGROUND TO THE RESEARCH ENTERPRISE

de Landsheere (1997) provides an historical account of educational research which looks at developments in distinct periods over the past 120 years and it is useful to consider the chronological account he outlines to better understand the evolution of educational research and its main characteristics (pp. 8-16):

pre-1900: Characterised by 'experimental psychology' in Europe and the United States that was concerned with children's language and imagination. The beginning of the 'child study' movement.

1900-1930s: A 'heyday of empiricism' which focused on rational management of instruction and saw the development of descriptive and inferential statistics.

1930s – 1950s: A loss of impetus in the strict scientific approach, due to philosophy taking precedence over science by valuing life experience over experimentation.

1960s - 1970s: A 'knowledge explosion' due to technological advances and the ability of affluent societies to fund research activities at unprecedented levels. Also, the beginnings of an epistemological debate in the social sciences, underwritten by the belief that rigid science could not accommodate the many perspectives of human behaviour and the subtleties of the social setting.

1980s - 1990s: An acknowledgement that no one research paradigm can claim hegemony in educational research and a softening of the hardline distinction between 'quantitative' and 'qualitative' methods. Also, recognition that no one research method can provide knowledge of the *true nature of phenomena*.

Despite the above indicating periods in which other ways of looking at the world achieved some prominence, much of the research enterprise in the twentieth century was dominated by a faith in broadly technocratic social science concepts of modernity, drawn largely from a functionalism[5] which had its origins in the French Enlightenment and perhaps as early as the birth of modern science in the seventeenth century. As Welch (1999) argues, however, the past few decades have seen a 'collapse of certainty' with regard to this theoretical way of seeing the world. It has been "breaking up, revealing an increasing fragmentation of purpose, and perhaps failure of vision" (pp. 25-26).

[5] Sometimes also called 'structural functionalism'.

Tracing the Fragmentation of the Research Enterprise and the 'Collapse of Certainty'

To describe the main characteristics of the present traditions of quantitative and qualitative approaches to research is, necessarily, to give an account of its history. That history should reveal not merely the characteristics but also the origins of the divide and the flaws in each approach. Given the progress that the natural sciences made throughout the nineteenth century, it is not surprising that the framework of the social sciences which began to emerge around that time was influenced by, and consequently shaped on, scientific models. For example, Comte's 'sociology' and John Stuart Mill's 'experimental psychology' inherited strongly the idea of 'functionalism' that was derived from the methods of the natural sciences, in particular physics. Herein, too, lay the origins of modern educational research (de Landsheere 1997, p. 9; Husén 1997, p. 17; Welch 1999, p. 28). Applied to education, functionalism describes the role of scientific reason, where the use of a supposedly value-free social science would allow researchers to simply seek out and present the facts, avoiding questions of ethics or any moral dimensions of knowledge they advanced. One can see how this approach would have seemed desirable to civic and political leaders of the time in terms of giving a developing society a precise method by which it could pursue its social-cum-national goals. It was an opportunity to deliver a precise future by means of positivistic investigations through scientific methodology, which had the capacity to herald in the same golden age of progress, and discovery that was the hallmark of natural sciences at the time (Welch 1999, pp. 27-28).

As de Landsheere reports, the first few decades in the 1900s were giddy days of empiricism, where educational research, armed with a framework which was steeped in the tradition of the natural sciences, focused on rational management of instruction and advanced the development of new curricula, psychological testing, administrative surveys, and normative achievement surveys. In addition, researchers were managing their data with the already established methods of descriptive statistics, as well as using the inferential statistical concepts that developed throughout the 1920s and 1930s (p. 8). In our view the use of the term natural sciences conceals the overriding importance attached to the methodology of physics and, in particular, classical mechanics.

The period from the 1930s to the late 1950s gave us perhaps the first indications of a social environment, which would lead to questioning the suitability of natural science's framework of subject-object dualism for research into the human condition. The catalyst for this was a compounding of the Western economic crisis of 1930s and the impact of World War II on countries across the globe. The 1929 U.S. Stock Market crash brought about a massive economic slump that spread to virtually the entire industrialised world and lasted for a decade. Understandably, research funds became very scarce. Then, in 1939, the world descended into World War II; the most bloody conflagration in human history. Whilst educational research continued in some countries, it more or less ground to a halt in others, either as a result of a redirection of focus on the war effort or, in the case of Germany and Russia, because "freedom of research was not acceptable to dictators" (de Landsheere 1997, p. 13). Despite a continued interest in educational research based on natural science, de Landsheere reports that two factors in addition to the world economic crisis and WWII led to the strict scientific approach losing some impetus to a more philosophically oriented and innovative progressivism, (a) the atomistic character of most educational research, and (b) the appeal of combining empirical research with social and political philosophy, which merged free enterprise and a liberal spirit with humanistic socialism (p. 8).

Of interest, the emergence of 'critical theory', or neo-Marxism, in this period is linked to a number of scholars in Germany who not only recognised the injustices and social distortions that could be perpetrated by political regimes (themselves having had first-hand experience with this under Hitler), but also rejected the all-pervading influence of positivism which championed 'instrumental rationality' and viewed all practical problems as technical issues (Bosetti et al 1989, p. 3). The 'Frankfurt School' held that the popular appeal to 'scientism' created an illusion of an 'objective reality' over which the individual has no control, and hence a diminished capacity to reflect on and change their own situation. By advocating the Weberian principle that *ends, means, values and facts* are conceptually separable, the scholars maintained that scientism led to a *bureaucratic rationality* which promoted a *false consciousness* whereby the prevailing social mechanisms bind some groups to accept irrational and distorted ideas about their social reality (Carr & Kemmis 1986, pp. 96-130; Rizvi 1986, p. 3).

The economic and technological developments in the period after World War II culminated with an explosion of knowledge by the 1960s and 1970s, fuelled by the availability of fiscal and technical (including computer) resources. The dynamism of the post-war era was palpable and educational research, particularly in the United States, was funded at unprecedented levels by the public and private sectors (de Landsheere 1997, p. 8, 13-14). It is ironic that this apparent time of plenty also crystallised the challenges to natural science as the best way to conduct research into human activity and behaviour. As Keeves (1997) reports, the increased funding for educational research provided scholars in the social sciences with opportunities to work collaboratively on education issues with historians and philosophers (p. 4). Whilst this was important in introducing new perspectives and methods into fields such as sociology, anthropology, politics, history, and philosophy, what was happening in broader society was paving the way for what Welch (1999) calls 'fin de siècle fractures' of a rationalist ideology of perfectibility (p. 35).

By the late 1960s, there was a perception of a 'cultural crisis' in industrialised societies which was derived from a disappointment that neither science and technology, nor middle-class values, were solving problems in the sense of providing general peace, wealth, and happiness (de Landsheere 1997, pp. 13-14). More concrete examples are provided by Welch (1999):

> "The oil crisis of the 1970s, as well as periods of intermittent recession thereafter, led to the advent of mass unemployment, especially among the young … The widening gap between rich and poor (both within and among countries) and the increasing deregulation of many economies evidenced a more general decline in government activity and intervention in social and economic affairs. In the social sciences, the confident certitudes of earlier decades were falling increasingly into disarray." (p. 35)

This despair of this period reflected a time when not only was the nature and use of scientific method questioned in the social sciences, but so were the foundations of science itself. According to Burns (1999), "Science ... lost its aura of eliteness and sacredness, which in the past has prevented researchers from questioning its assumptions" (p. 12). The technology and knowledge that had promised so much for humanity, looked unable to solve the profound social and environmental problems arising out of our emerging modernity. Herein lies the seed for what would become the 'poststructuralist' thinking which would reject much of the modernist platform by challenging the assumption that the dictates of technology of reason would promote a more rational and more morally perfect world (Welch 1999, pp. 28-35). It should be said that it was not so much an attack on the *idea of science* per se; rather, it was a call for science to be accountable in terms of the logic of its methodology, its knowledge claims, and its apparent detachment from questions associated with value, ethics, and politics (Connole 2000, p. 18).

This is an account of one view of the collapse of scientific certainty. From the view of the theoretical physicist the collapse had occurred much earlier. In 1795 Karl Friedrich Gauss entered the university of Gottingen having already devised the gaussian curve. With that curve came the realisation that the error can never be taken out of observations. The achievement of twentieth century physics was to show emphatically that Gauss was right and there is no such thing as an exact picture of the world of our reality. In 1927 Heisenberg stated his principle of uncertainty and that uncertainty is mapped out by Planck's quantum (Bronowski 1973). It took some time for educational researchers to realise what had happened. The view of physics on which they had based their view of scientific research in education was passé. Society at large still does not appear to have grasped the idea.

There emerged two ways to research educational problems which Husén (1997) describes as separate approaches that are either 'functional-structural, objective-rational, goal-directed, manipulative, hierarchical, and technocratic' *or* 'interpretivist, humanistic, consensual, subjective, and collegial' (p. 19)[6]. Overall, it is fair to say that there has been a great deal of tension between the various methodologies to the point where the investigation of educational problems has suffered as a consequence, through schools of education being split, editorial policy of journals being influenced by research approaches, preferential funding for certain projects, and administrators and politicians belittling research in education (Keeves 1997, p. 1).

THE TRADITION OF THE QUANTITATIVE APPROACH

Quantitative approaches to research are often termed 'empirical'. Empiricism, however, is not a methodology as such but a thesis that knowledge is gained by observation of real events. As far back as Ancient Greece, scholars began to base important conclusions about nature on their observations at the expense of more widely held non-empirical conceptions such as mythology, religion, and appeals to authority. Indeed, many great advances in science facilitated by people such as Hippocrates, Copernicus, Galileo and Darwin are attributed to empiricism as a way of gaining knowledge (Graziano & Raulin 1996, p. 8). An empiricist believes that reality is unitary and it can be understood by empirical analytic enquiry, a method which collects data verified by the senses to form or test a theory (Connole 2000, p. 32; Doordan 2000, p. 57). With regard to social reality, the quantitative approach is *nomothetic* in assuming that it is objective and external to the individual (Burns 1999, p. 3).

Proponents of a quantitative research claim to use a 'scientific method' which has the characteristics of *control, operational definition, replication,* and *hypothesis testing* (Burns 1999, pp. 5-7). They would argue that their approach provides advantages such as;

1. Distancing the researcher from the object of enquiry in the sense that the enquirer is independent from those being researched. As such, the findings should not be influenced by the researcher (Doordan 1998, p. 140);
2. Demanding that the description of phenomena involves precise, unbiased recording of observations. There is an emphasis, therefore, on a research design which is based on measured, quantitative information which can be analysed by statistics to support or disprove claims;

[6] This paper considers the 'quantitative/qualitative' divide. It recognises the existence of 'critical theory' and 'postmodern / deconstructive' approaches, but collapses them into the qualitative framework for the purpose of the argument.

3. Eschewing relativist and metaphysical claims as well as seeking objectivity by removal of errors, values and biases (Connole 2000, p. 41; Doordan 1998, p. 140);
4. Maintaining an emphasis on replication, prediction and control, and
5. Being based on *philosophical determinism* which is the notion that events occur according to regular laws and, as such, human behaviour can be viewed as an outcome of antecedent environmental events (Connole 2000, p. 41).

At this stage it should be made clear that empirical or positivist investigations can occupy a variety of positions, ranging from 'hard empiricism' or logical positivism (which asserts that the scientific method is the *only way* to test knowledge claims) to 'soft empiricism' (which lends itself to empiricist and non-empiricist methods) (Connole 2000, p. 40).

THE TRADITION OF THE QUALITITATIVE APPROACH

Interpretivists (or internalists) maintain that we use constructs such as culture, social context and language to build our view of the world and that social reality is shaped through social interactions (Smith 1989, p. 74). Implicit in the interpretivist notion of existence is *intentionality* which refers to a state of 'being in the world' and infers an interdependence between 'thought' and 'lived experience', predicated on 'meaning' derived from 'social interactions' (Smith 2000, pp. 97-98). Because existence is viewed in this way, interpretivists are committed to an epistemology which embraces *social constructivism* and knowledge as 'meaning in context' made possible by social interaction (Williamson 2000, p. 30). Social knowledge is not, as empiricists (or externalists) claim, something which exists independently or external to us and waiting to be discovered by untainted sensory perception. Interpretivists say that you cannot have unmediated access to reality. Instead, they see the world as mind-dependent. As such, it appeals to philosophical notions of relativism, existentialism, and phenomenology.

Given these interpretivist tenets, it follows that their research interests will be concerned with people's beliefs, feelings and interpretations and how they make sense of their world through meaning (Williamson 2000, p. 31). The roots of this form of enquiry lie in *hermeneutics* which originally referred to the interpretation and understanding of scriptural texts. Gradually the term has come to be used in a wider context to include human actions, customs and social practices (Williamson 2000, p. 141). Interpretivism is an umbrella term normally associated with qualitative research methods[7] for evaluation in the social sciences. The research is field-focussed and natural settings are the direct source of data in which the researcher is the key instrument. The researchers are concerned with process rather than simply outcomes and the research is *idiographic* in the sense that it studies individual cases (or small groups) intensely. The data is descriptive, *thick, and rich*[8] and has traditionally been collected by means of words and pictures rather than numbers. Furthermore, data collection and analysis occur simultaneously and theories are developed inductively from the 'bottom up' by means of the evidence which is collected (Smith 2000, pp. 101-105; Bogdan & Knopp Biklen 1992, pp. 29-32). de Landsheere (1997) acknowledges a diversity of approaches employed by the humanistic research movement, e.g. anthropology, sociology, politics, history, linguistics, philosophy, and ethnomethodology (pp. 8 & 13). Moreover,

---

[7] It should be noted that some interpretivist approaches (e.g. *grounded theory*) do use quantitative techniques. In addition, we assert that there are qualitative approaches that lie outside of the 'interpretive umbrella'.

[8] Data which goes beyond fact and surface appearance by presenting detail, context, emotion, feeling, spatial and temporal character, voice, meaning and interrelationships (Smith 2000, p. 198).

interpretive approaches can be generally classified by their purpose, orientation and report form[9](Smith 2000, p. 118):

Interpretive orientation	Explanatory	Descriptive	Expressive
Report form	Presentational	Explication of issues	Representational
Research purpose	Theory generation	Portraying complexity	Eliciting ethical reasoning
Research methodologies	Grounded theory	Qualitative case study	Narrative enquiry

Ethnomethodology ⟵ ⟶ *Phenomenology*

THE POSITIONING OF EDUCATIONAL RESEARCH

Now that the two main approaches to research have been made explicit in terms of how their proponents would present them, it is timely to consider how each is situated in the research enterprise in general. To assist in this understanding, it is useful to consider the model put forward by Popper and Eccles (see Figure 1 next page) that posits that there are three different worlds involved in human enquiry (Keeves 1997, p. 1-3):

World 1 (the Real World) - comprised of entities, including natural physical objects, people, and human constructions and activities;

World 2 (the Learner's Mind) - subjective experiences, mental states, conscious thought and psychological dispositions, unconscious states of mind, human wisdom; and

World 3 (the Body of Knowledge) - an objective world comprised of the corporate body of propositional knowledge concerned with causal explanations. Includes human endeavours of art, music and literature which are part of the world of shared knowledge.

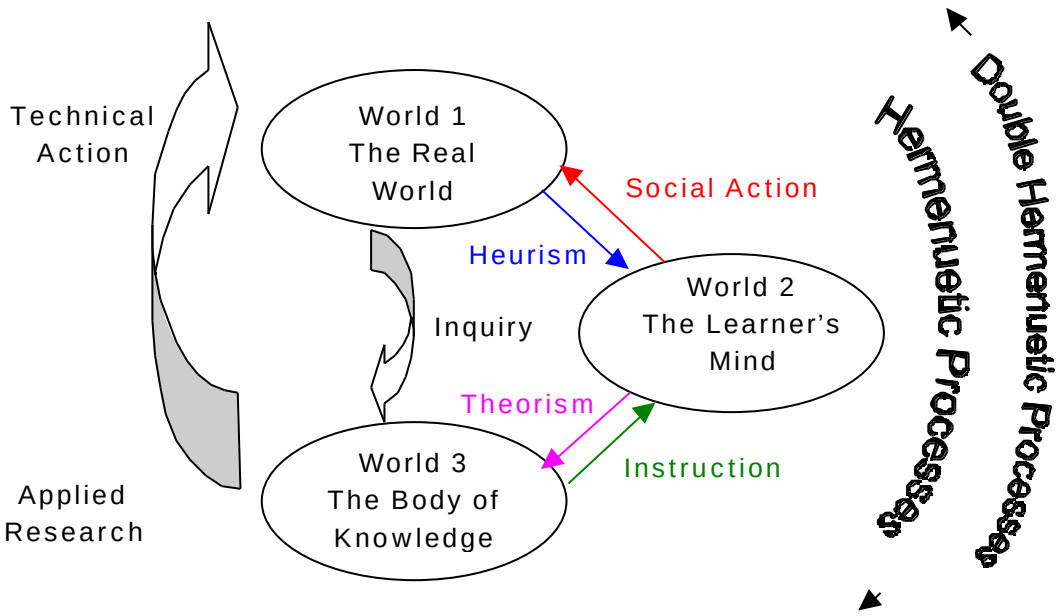


Figure 1: The Nature and Processes of Educational and Social Enquiry[10]

[9] The diagram is not meant to be exhaustive. It merely lists some of the commonly encountered interpretivist approaches and how they are generally classified by virtue of their characteristics. Also, it is important to recognise that a particular research design may utilise a 'hybrid' of approaches to achieve its goals.

The model conveniently separates our one, complex world into three and provides an insight into cognitive processes associated with learning and enquiry. In the same way that each of us can look at our 'one world' from a certain perspective, we can now dissect it and transpose this view onto the three worlds in the model. This is useful for understanding how the various research approaches are traditionally distinguished from each other, particularly with regard to the way that they view 'reality' and 'knowledge' as they apply to Worlds 1 and 3 (see Table 1 also).

All research approaches claim to investigate World 1, the Real World, but it is the way they view this world and what they feel is generated by the enquiry which differentiates them[11]. Both the quantitative and qualitative researchers would claim that their investigations of World 1 enable them to make contributions to World 3. It is clear, however, that they see World 1 in fundamentally different ways. For quantitative researchers, World 1 (the Real World) is independent from what the human mind thinks it is. For qualitative researchers, however, human thinking makes World 1'what it is'. Both camps would recognise World 2 although, again, from different perspectives. From yet another perspective, the criticalists and postmodernists maintain that World 1 and World 2 are distorted realities and that World 3 is made false by value-laden knowledge claims which are ideologically driven and which promote oppression. They believe, therefore, that traditional empirical and interpretative approaches, by lacking a critique of the worlds they investigate, do little to change social reality and, therefore, assist hegemonic powers to perpetuate conditions which dominate and oppress.

Paul and Marfo (2001) present what is essentially a political argument for the recognition of a paradigm shift in educational research. Pointing to the hegemony of quantitative methods in educational research for most of the 20th century they assert that there is currently, and has been for two decades, an increasing awareness and use of qualitative methods in educational research. Quantitative research they link to logical empiricism and qualitative research they link to an awareness of philosophical, moral and political values and their effect on the content and results of educational research. They then argue for the expansion of researcher's knowledge of the history, philosophy, sociology and ethics of inquiry so that educational research will cease to be pursued solely in a quantitative vein. But at the same time they point to a paradigm shift in the way that science is conceived such that the shift is:

> "…not about a counterforce; it is more about the collapse of the intellectual framework of empiricist science. The issue is the normative understandings that guide the work of social scientists. The "new science" is not an add-on but rather a change in the way we think about, conduct, share, and defend research." (Paul and Marfo 2001, p. 527)

They argue for the recognition of multiple paradigms of inquiry with no one paradigm establishing a hegemony.

Four years earlier Walker and Evers (1997), in considering the same issue. had distinguished three theses: the oppositional diversity thesis; the complementary diversity thesis; and the unity thesis (pp. 22-30). They argue that there is a quantitative/qualitative debate in which two fundamental paradigms are distinguished. The quantitative paradigm is linked to scientific methodology and the use of mathematics and measurement. The qualitative paradigm is linked to the interpretation of human action by attempting to understand it. They refer to Husén and quote with approval his summation of the distinction;

---

[10] Diagram adapted from its presentation in Keeves 1997, p. 2

[11] It is also logically possible to investigate Worlds 2 and 3 as well the connections between them, although this would probably be done via their existence as World 1 phenomena.

> "The twentieth century has seen the conflict between two main paradigms employed in researching educational problems. The one is modelled on the natural sciences with an emphasis on empirical quantifiable observations which lend themselves to analyses by means of mathematical tools. The task of research is to establish causal relationships, to explain (Erklären). The other paradigm is derived from the humanities with an emphasis on holistic and qualitative information and interpretive approaches (Verstehen)." (Husén 1988, p. 17)

The three theses are ways of conceiving of the quantitative/qualitative distinction and its effects on educational research;

1. The oppositional diversity (OD) thesis revolves around a conflict between quantitative and qualitative approaches which is a conflict between a paradigm that presupposes a mind-independent reality where research is an attempt to produce something that corresponds to reality and a paradigm that presupposes that there is no mind-independent reality and research is aimed at understanding the constructs that human beings produce and which guide their actions. The weaker version would confine itself to making such claims about social reality and emphasise that educational research is research about the social reality of human beings. The paradigms are incommensurable.
2. The complementary diversity (CD) thesis supports a notion of paradigmatic pluralism. Different paradigms have their place in researching the complexity of the education of human beings. No one paradigm is capable of doing the job. The paradigms are incommensurable.
3. The unity (U) thesis is the thesis that there are not distinct, incommensurable paradigms but paradigms that can work together or in competition to the benefit of the research. They argue for a 'touchstone' which enables a researcher to decide between approaches.

The OD and CD theses both presuppose that the paradigms are incommensurable. The U thesis argues that they are not. The OD and CD theses involve the idea of distinct paradigms existing within distinct traditions. The U thesis is the view that distinct paradigms whether or not in distinct traditions can work together in an enquiry into a common problem. The description of the qualitative paradigm is a restricted view of that paradigm to which we shall return later. We would also argue that the notion of a 'touchstone' is a Kantian myth which has been sought for over two hundred years without success.

Walker and Evers (1997) commence by saying that;

> "The major epistemological question here is whether these distinctions are associated with different ways of knowing or forms of knowledge, which partition educational research so that research traditions, for example, turn out to be radically distinct epistemologically, each having its own theories and rules of justification, meaning, and truth. If so, the next question is whether findings produced by the different traditions can be rationally integrated, rendered coherent, or even compared. For this to be possible, for traditions to be commensurable, there will have to be some shared concepts and standards of justification, meaning, and truth; some epistemological touchstone. If, however, the traditions are so fundamentally disparate that any choice between them in educational research is arbitrary or the result of nonrational commitment – an act of faith – there is no touchstone. The research traditions are incommensurable." (p. 22)

We shall use the term traditions as indicating distinct ways of looking at the world of our experience in that they are founded on a distinct set of ontological and epistemological beliefs. This set includes basic approaches to rationality. Paradigms, or in Walker and Evers term, theories, exist within traditions. We shall then argue that traditions and the paradigms within them are not incommensurable and can be rationally compared and evaluated thus leading to the conclusion that quantitative and qualitative approaches can be complementary. We reject the touchstone argument, it is in our view a consequence of the Kantian attempt to find axioms which any rational person must accept. This attempt has always been a failure.

Within the natural sciences there have been a succession of paradigms. For instance, in physics, classical mechanics was followed by Einstein's view of the world. One paradigm succeeded another. These paradigms are ways of conceiving and explaining reality and while they can be very different they all arise within the same tradition of enquiry, a tradition that has a particular history and presupposes the possibility of an empirical explanation of reality without reference to any norms. In contrast to this there is the paradigm that gave way to classical mechanics, medieval Aristotelian mechanics. This paradigm was conceived within a different tradition which had at its base the conception that all things had a final end. Aristotelian mechanics had a teleological base in that it arose within a tradition that viewed the world and its inhabitants through a teleological perspective. The practice of physics within a tradition is pursued from the presuppositions of that tradition. This directs the practice towards particular explanations and admits particular paradigms. So in Aristotelian mechanics things fell towards the earth because they were seeking their preferred state – a teleological explanation. In classical mechanics things fall towards the earth due to gravitational attraction – an explanation compatible with presuppositions that recognise an empirical/valuative distinction. All traditions arise and live within a culture and their history in the beginning is a history of that culture. The narrative of the history of a tradition can, from its original beginnings, encompass many changes and modifications. The Aristotelian tradition is still with us having passed through medieval times and the Galilean revolution. It has changed with the passage of time. Aristotelian mechanics is the mechanics of a tradition born in an Athenian culture of a particular time. The tradition was changed and modified as time passed but even in the 13th century it retained its teleological approach. A different tradition had its seeds in Galileo's work and was formed by the Enlightenment. Any given tradition may thus contain a number of paradigms which in turn may change and be modified and be discarded without destroying the tradition. Traditions are distinguished by their particular set of ontological and epistemological commitments.

There is a radical difference between competing paradigms within a tradition and competing paradigms from different traditions. When Boltzmann set out the anomalies that arose in attempting to give an account of thermal energy within the paradigm of classical mechanics there existed a crisis in classical mechanics. The crisis was resolved by Bohr's theory of the atom. But the resolution was more than that. It modified the paradigm. The point is that within a tradition there can be conflict and argument and change. In fact it is better to say that change is necessary within a tradition if it is to continue to be a live force. Where anomalies exist, which they must, then the epistemological struggle is one to relieve the paradigm of the incoherencies and produce a new or modified paradigm which at one and the same time solves the incoherencies and gives an account of why they were bound to occur. In this way a paradigm within a tradition has a continuous history and the narration of that history is an account of the enquiry, its success and its change over time. The conflict between paradigms within a tradition is a conflict fought and resolved using common epistemological and ontological standards, the standards of that particular tradition.

The situation of competing paradigms from different traditions is one that has fuelled the belief that the paradigms are incommensurable. Each paradigm within each tradition will, to the extent that the tradition is coherent, exist and be evaluated according to the internal standards of the tradition. If these standards are incompatible as between traditions, which they must be if the traditions and paradigms are to be distinguished, then each tradition and paradigm is adjudged incoherent by the standards of the other. One response to this is to seek a touchstone. The incommensurability is derived from the differing epistemological and ontological standards of the different traditions within which the paradigms exist. To resolve

the incommensurability a source of epistemological and ontological standards are sought which are outside of and independent of the incommensurable traditions.

Another response is to deny that rational debate between competing traditions is possible, everything is relative to the tradition within which it exists. This is then enlarged to the usual relativist thesis concerning cultures and knowledge. If both traditions are coherent within themselves but incompatible with other traditions then no tradition can make any successful claim to truth. On this basis an Aristotelian approach is incompatible with a Newtonian approach and, because no external independent touchstone is possible, then, no resolution is possible. An Aristotelian teleological approach to natural science may be internally consistent but incompatible with a Newtonian empirical approach but there can be no possibility of a decision as to which is more reasonable. We suggest that this is, in part, the thinking that lies behind the suggestion of multiple paradigms by Paul and Marfo and by the arguments of Walker and Evers.

We suggest that it is a mistake to think that distinct paradigms existing within distinct traditions are incommensurable. Consider first the assessment within a tradition of an existing paradigm facing attack from within as classical mechanics was from Boltzmann. The first step in assessing a paradigm is always to attempt to show that by the internal standards of the tradition the paradigm is incoherent. This Boltzmann did. The second step is to attempt to show that a new proposal solves the incoherence and gives an account of why it was bound to occur. This Bohr did. Note that the second step will involve giving a narrative of the paradigm for it came into existence originally to solve certain problems which had presented themselves. A history of a paradigm is, on this view, an essential part of understanding the paradigm, its weaknesses and whether or not a solution has presented itself. Note also that the step involves showing why the problem was bound to occur. This goes to an assessment of the weakness of the paradigm and provides a justification for seeking a solution. A third step is not essential but would consist in starting to consider what incoherencies are bound to occur within the new proposal. The presupposition here is that while the final end of the enquiry may be the truth, the pursuit of the enquiry will yield only steps towards that final end and these steps will necessarily contain flaws which hopefully will be corrected (Einstein and Infeld 1938). This serves to direct the future of the enquiry. Why is it thought that this methodology cannot be followed when considering paradigms within competing traditions? It is essentially a rational methodology and should be applicable. However, let us first apply the methodology to quantitative research in education.

Educational research is research about human beings. In the 17^{th} and 18^{th} centuries the empiricist explanation of a human action was based on a search for the physiological mechanisms. Human action was conceived of in mechanistic terms and it was supposed that the study of human action would yield universal laws in much the same manner that Newton's study of the world had yielded the laws of dynamics. In such a situation human action would become predictable much as the orbit of the planets and comets, both known and unknown, were predictable.

Kant accepted the incompatibility between any account of human action couched in mechanistic terms and the language of morals. Moral language was concerned with human action based on maxims and rules. The science of human action was based on the search for and exposition of the mechanical causes of the ways that human beings behaved. Finding the two incompatible Kant asserted that moral language is inexplicable in scientific terms (Benn and Peters, 1959, p.47). There is a gap between science and values which cannot be bridged because there is no possibility of explaining the inexplicable. So concepts central to morality

such as intention, purpose, virtue, and good which we ascribe to human action become unintelligible to those pursuing scientific research into human action. The gap between 'ought' and 'is' becomes an accepted part of the intellectual landscape and the acknowledgement of its existence is seen as an achievement of modern, that is, Enlightenment, thought.

> "This distinction between normative rules and scientific laws, which is here regarded as basic to our understanding of society, was made explicit comparatively late in the history of thought – probably in Europe in the eighteenth century." (Benn and Peters 1959, p. 16)

To the present day it has been accepted that a mechanistic explanation of human action, if it is to remain true to its attempt to explain human action by quantifying it and seeking universal laws akin to those of natural science, must find human action inexplicable as described in terms of intention and purpose. Quine has argued that any attempt to study human action scientifically must eradicate all mention or allusion to intention and purpose. This is the origin of quantitative research and for many it has not and should not deny its origins.. The task is to emulate the physics of classical mechanics.

Is the quantitative research paradigm incoherent? Consider a stranger drinking a liquid from a container. We can describe the mechanics of the situation. If we have sufficient physiology we can describe what is happening beneath the exterior of the body that we see before us. We can give the circumstance a cultural context by recognising that the container is a wine glass and that the liquid is red. But while we can describe the mechanics and by looking at the cultural context we can guess that the stranger is drinking red wine we cannot with any surety give a description of the human action which is taking place. The act of drinking the red wine may be described as:

> "…an act of self-indulgence, an expression of politeness, a proof of alcoholism, a manifestation of loyalty, a gesture of despair, an attempt at suicide, the performance of a social rite, a religious communication, an attempt to summon up one's courage, an attempt to seduce or corrupt another person, the sealing of a bargain, a display of professional expertise…" (Ayer 1964, p. 7)

The point being that the stranger is the final authority on what he or she is doing. How he or she sees their behaviour is what makes the action the action that it is. The mechanics are the same; the action differs. In the same way measurement of human behaviour may capture the mechanics of the circumstance but misses the humanity of the behaviour. This is the problem that quantitative research of human action must face. It is a problem that arises from following the empirical/normative distinction in the sphere of human action.

Consider a metal bob attached to a string which in turn is attached to a beam. The bob is swinging to and fro. If we ignore the air pressure, the friction at various points, the elasticity of the string and keep the arc through which the bob swings below 30° then, by using mathematics we can come to the formula "$T = 2\pi\sqrt{l/g}$" where 'T' = period of the bob, 'l' = length of the system and 'g' = gravitational constant. This is a curious achievement. All mathematics can be represented as a system of exact concepts. An exact concept is one that does not permit of the neutral case.

> "…that is, a case in which both the assigning of the concept and a refusal to assign conform to the concept-governing rules. To put it another way, a circle is either a circle or it is not … there is no such thing as a borderline case." (Gibbons 1979, p. 323)

On the other hand the physical system of the bob can only be given an inexact physical description for all empirical concepts are inexact. The sets of exact concepts and inexact concepts are mutually exclusive and the consequence of this is that in order to attempt a mathematical description of the a swinging bob, it is necessary to idealise the situation. That is, the situation must be made an exact situation so that mathematical concepts have purchase. As a result the formula of the simple pendulum "$T = 2\pi\sqrt{l/g}$" is divorced from

reality. This was precisely Galileo's point when his critics pointed out that no simple pendulum conforms to a standard swinging pattern. The key to physics, and to any science that uses mathematics, is to understand that an ideal description of the system must be created so that mathematics can do its work. The problem is then to apply any mathematical results to the real world. There has to be something similar to the process by which the ideal situation was created but in reverse. Educational research in its quantitative guise has sought to use mathematics and it faces the same problem that physics and other natural sciences face. It faces the same problem which anybody faces if they intend to apply the results of the mathematics to the real world. A good example of the way in which the problem is actually overcome is the use of 'fudge factors' by engineers to compensate for discrepancies between theoretical engineering ideas and their application in the real world.

Quantitative research was bound to face problems arising from the nature of human beings and the nature of mathematics. As a result, it becomes incoherent.

Qualitative research, as defined by Walker and Evers, comes under attack from all the arguments that have been used against constructivism and relativism. It is beyond the scope of this paper to rehearse those arguments; they have been expounded in detail by Kukla (Kukla 2000) and there is sufficient evidence to argue that it is reasonable to consider that qualitative research as defined by Walker and Evers is incoherent.

There is a further argument that is relevant. The point of educational research is not to enlighten us about the particularities of an individual but to enlighten us about groups. Research in a qualitative mode which eschews any quantitative approach necessarily confines itself to the individual and cannot extend its findings to the group. It is the same problem as that of applying mathematical formula to a 'real world' system. In order for the findings about a unique individual to be generalised across a group, the group must be idealised in the sense that certain features are used, others ignored. The situation is analogous to the pendulum and the problem is the same.

The U thesis argues that the quantitative and qualitative paradigms are commensurable. This can be taken to mean that either it is possible to compare paradigms from different traditions or that both the quantitative and qualitative paradigms exist within the same tradition. We would argue that the quantitative and qualitative paradigms as defined by Walker and Evers exist within different traditions and therefore they are committed in the U thesis to the possibility of comparison between paradigms from different traditions. It is unclear how they expect this to occur. They also appear committed to the notion that the quantitative and qualitative paradigms can work together, in a sense combining their strengths. An example of the way that this may be taken to occur has come about with the advent of powerful computers and specific software.

Over the past thirty or so years, one way that people have distinguished the two approaches is by designating *quantitative research* as 'number-crunching' which seeks explanation by reducing reality to its component parts and representing it numerically as formulae or establishing its place in the world by means of descriptive or inferential statistics. It would not be drawing a long bow to suggest that this characteristic has often been a way in which researchers from both approaches have given each other a hard time. The qualitative researcher says of the quantitative researcher "they're just number-crunchers", to which comes the retort "your lack of such activity confirms my suspicion that your research is not good research". With the advent of more compact, efficient, powerful, accessible, and transportable computers, however, the time of this superficial distinction can be said to be well and truly over.

> "Since the 1960s, the computer has become the daily companion of the researcher. For the first time in the history of humankind, the amount and complexity of calculation are no longer a major problem. Already existing statistical techniques, like multiple regression analysis, factor analysis, multivariate analysis of variance, that previously too onerous for desk calculation, suddenly became accessible in a few moments." (de Landsheere 1997, p. 14)

The important point here is that not only is this seen as benefiting quantitative researchers, but many qualitative researchers are beginning to sort and code data for computer analysis (Burns 1999, p.14). The most commonly used packages include ez-text, ETHNOGRAPH, Atlas.ti and NUD*IST, all of which offer a qualitative content analysis package for verbal material. Another more recent software application, NVivo, takes qualitative inquiry beyond coding and retrieval. It integrates coding, links data to external multimedia data, websites or internal documents or concepts, stores ideas flexibly in annotations and rich text memos that can be coded, linked and searched, manages rich data, and allows information in tables to be imported from or exported to statistics packages (QSR 2002).

It is interesting to observe mixed reactions to the emergence of computer assisted qualitative data analysis software (CAQDAS). Some researchers have embraced the opportunity that automation provides in terms of (i) speeding up and enlivening the coding process, (ii) providing a more complex way of looking at the relationships in the data, (iii) providing a formal structure for writing and storing memos to develop the analysis; and (iv) aiding more conceptual and theoretical thinking about the data. Others, however, see the introduction of software as the 'the dark side of the technological advance' and are concerned that it will (i) distance people from their data, (ii) lead to qualitative data being analysed quantitatively, (iii) lead to increasing homogeneity in methods of data analysis, and (iv) that it might become a 'monster' and hijack the analysis (Barry 1998). There is another problem. Does the coding of qualitative results purport to treat valuative terms mathematically in defiance of the is/ought distinction, or does the coding of qualitative results in mathematical terms proceed by ignoring the valuative elements and thus preserving the is/ought distinction? Has Hume and the Enlightenment been abandoned or does it continue to reign (Couvalis, 1997, p. 2)? Whichever answer is given, wherein lies the justification? We suggest that the users of such computer software must confront the ontological and epistemological issues involved in their use.

COMPARING ACROSS TRADITIONS

We propose to attempt an argument to show that paradigms from different traditions can be compared rationally and also that there is a paradigm of qualitative research which has so far received no mention and which is in current use in some natural sciences.

A paradigm exists within a tradition and a tradition is, as the name suggests, something that has its origins in a particular society and the beliefs of that society. The tradition will, while it remains alive, grow and change throughout its history. The history of a tradition is an account of the existence, change and coherence of the tradition. Consider Galileo. Aristotle, that philosopher of common-sense, dominated thinking about mechanics for over 1800 years. It was Aristotle's mechanics that Galileo studied as a student at Pisa and his first work as a scientist was to annotate Aristotle's *De Caelo.* There are some aspects of Aristotle's mechanics to which we wish to draw attention and compare with Newtonian mechanics.

Firstly, for Aristotle, all motion required a moving force. This force could be, in the case of heavy bodies, the attraction towards its natural place, the centre of the earth, or a force applied to the body such as a push. Inertia was a state of rest and all motion had to be accounted for by reference to a force. This accords well with our everyday experience, for an

object does rest unless it is pushed and will stop moving if we cease pushing. Motion is a change of position. And it was in giving an explanation of such observations that Aristotle was led to say;

> "Everything that is in motion must be moved by something. For if it has not the source of its motion in itself it is evident that it is moved by something other than itself." (Aristotle 1984, 241b, pp. 34-35)

At the centre of Aristotelian dynamics is the idea that bodies move against constant resistance; inertia is defined as a state of rest. Our language contains within it the Aristotelian idea of inertia. The dictionary describes 'inert' as having no inherent power of action, motion, or resistance and 'inertia' as inactivity.

This is in contrast to Newtonian mechanics. The axioms or laws which appear in the *Principia* can be stated as:

1. Every body continues in its state of rest, or of uniform motion in a straight line, unless it is compelled to change that state by force or forces acting upon it;
2. The change of momentum of the body is proportional to the force acting upon it; and is made in the direction of the straight line in which the force is directed;
3. To every action there is an equal and opposite reaction.

The basic principle of Newtonian mechanics lies in the first law. It is a law of inertia and is in direct contrast to that of Aristotle's. It was prefigured by Galileo with his axiom that a moving body on a frictionless level surface would continue to move with uniform motion to the limits of the plane. Galileo's notion envisages motion around the plane of the earth, Newton's motion was in a straight, Euclidean line. For Aristotle, space is finite and limited to that space within the sphere drawn by the fixed stars. Outside that sphere there is nothing. For Newton, on the other hand, space is Euclidean. It is infinite.

To grasp Newtonian inertia we must imagine an idealised system. Aristotle demanded no such exercise of the imagination. Consequently, any student with sufficient command of the English language to use either 'inert' or 'inertia' is confronted with a problem when confronted by Newtonian dynamics for they will tend to speak an Aristotelian language.

Newton's second law is usually expressed as $F = m.\ dv/dt$ where 'F' = force, 'm' = mass, and 'dv/dt' = rate of change of velocity with respect to time (acceleration). A more common form of expression is $F = m.a$ where 'a' = acceleration. Again, in contrast with Aristotle, we are asked to imagine an idealised system. Our everyday experience is of heavier bodies falling faster and there is never an absence of resistance.

Newton's third law is again a case of an ideal system in which there not only can be action and reaction in contact but also at a distance. In contrast, Aristotle's approach was to consider that there could not be action at a distance and that all action which produces a reaction must be by way of contact. Nor was the reaction considered equal and opposite but dependent upon the circumstances of occasion.

Aside from the detail of the laws, Newtonian mechanics demands that we see the world with which we are concerned in an idealised way. This approach was adopted by Galileo in his work on the pendulum and it enabled mathematics to be applied to physical problems.

Aristotle did not conceive of his mechanics as applying to an ideal system but to the real world. He is thus limited in the use to which he can put mathematics (should he have wished to do so).

Clearly, the learning of Newtonian mechanics by students poses problems. The main problems are that students are asked to put to one side their everyday notions, that is, their Aristotelian notions, and enter via their imagination into a different way of seeing the physical world which they inhabit. It might be thought that this is merely the problem of learning another approach to put alongside their previous commonsensical approach. Students, as people, are accustomed by the time that they reach adolescence to wearing different hats for different occasions; this is just one more hat. And yet the anecdotal evidence and the evidence in the literature suggests that students find this very difficult at times.

It is evident that science teachers are confronted with a problem. The world that their students observe is an Aristotelian world. The problem for the teachers is to get the students to change their point of view and look at the world as an idealised system in order to comprehend modern physics. This problem has been commented on before (Ebison 1993, Lombardi 1999) with little more than injunctions for teachers to be aware of the problem.

The problem is not insoluble; the solution lies in 'translation' from one tradition to another. And this, we would suggest, is the solution to alleged incommensurability. It is similar to the problem confronting language teachers when they attempt to get pupils to learn a second language. It is possible to learn a second language in the sense that it enables a person to get around a country and a culture without too much fuss but with little grasp or contact with that country and culture. If there is to be a worthwhile contact with that country and culture then the second language must become as a first language. Only then is it possible to understand and partake of the beliefs and attitudes of that culture. The same is true of different traditions. The scientist immersed in classical mechanics must learn the language of the Aristotelian tradition of mechanics if they are to understand the coherence and incoherence of that paradigm within a particular tradition. That is the first step. The second step is to ask whether or not the Aristotelian tradition brings with it insights which are superior or inferior to those from the tradition within which classical mechanics is embedded. In this way it is possible to compare and contrast paradigms within traditions rationally. Certainly there may be elements in a tradition and its paradigms which are incommensurable just as it is impossible to translate certain phrases from one language to another but, accepting this, it is still possible to make significant comparisons and assessments. For instance, the Italian 'simpatico' has no direct translation into English but can be understood by giving a lengthy explanation which would include references to Italian culture and belief.

Can the Aristotelian view of the change of motion of a body be expressed in the language of classical mechanics with sufficient exactness to enable a comparison and evaluation? Newton's second law is , $F = m.\ dv/dt$, or its more common form of expression, $F = m.a$. In the terminology of classical mechanics, it has been said that for Aristotle, $v = k\ F/R$ (where 'F' = force, 'k' = a constant, 'v' = velocity, and 'R' = resistance (Stinner 1994, p. 78). We would argue that the translation is sufficient for the purpose of a rational comparison. And this is a translation between traditions.

Quantitative and qualitative research, as normally conceived, are different paradigms within different traditions. To compare their merits the researcher in one paradigm must understand the other as well as their own. Throughout the argument so far the science that grew with the Enlightenment has been presented in terms of the idealised world of classical mechanics. This emphasised a break with the Aristotelian tradition. However one paradigm from the Aristotelian tradition was not discarded; that of the teleological paradigm. Both chemists and biologists have continued to think and pursue their enquiries in terms of 'natural kinds'.

Quantitative research has historically taken physics as its exemplar and this has in great measure produced the quantitative/qualitative distinction and divide. We would argue that the teleological paradigm as it occurs in chemistry and biology is worthy of consideration. It offers the opportunity to combine in a significant and complementary fashion, both quantitative and qualitative research. We have also argued that quantitative and qualitative research paradigms are commensurable. If so, then they can work together.

Keeves (1997) suggests that investigative approaches to educational research should embrace a coherent approach in terms of *methods* and *outcomes.* With regard to the former, "the methods employed in educational enquiry should ... be influenced by the nature of the problems being considered" (p. 6). With regard to methods, one should recognise the situational strengths and weaknesses of particular investigative instruments. Slade (2002) concurs and states that is best to remain open to employing a diversity of methods as the research problem demands and doing so exemplifies the logic, pragmatism, and importance of remaining flexible to meet the emerging needs of the investigation at hand (p. 98). These views entail commensurability. It is acceptable that *methods* from both the scientific and humanistic approaches can indeed be complementary in any given investigation. We have tried to show what questions must be asked if this is attempted. Further, it is logically possible for different methodologies, i.e. quantitative and qualitative, to work collaboratively in social science and educational investigations. Husén (1997) makes this clear by use of an example of research from teaching which 'mixes paradigms' constructively to demonstrate how human behaviour in a single classroom can be interpreted in a national or international context (by quantitative means) as well as the classroom itself as a unique phenomenon (by qualitative means) (p. 20). As de Landsheere (1997) suggests of research throughout the 1990s, it is no longer either-or, but both as "the scientific approach is seen to be complementary to the anthropological, historical, phenomenological, or humanistic approach" (p. 9). In this sense, the idea of a 'unity of purpose' in educational research can be supported in terms of respecting the capacity of quantitative and qualitative paradigms to add clarity to understanding human behaviour.

Keeves and Slade, go a step further, however, and argue that a 'unified' or 'coherentist' approach to research not only implies a complementary association between quantitative and qualitative methods and methodologies, but that the traditional separation of these research methodologies into different camps, often antagonistic to each other, is fallacious on epistemological grounds. We concur for the reasons developed throughout this paper.

REFERENCES

Aristotle. (1984). Physics. In Barnes, J. (Ed.), *The Complete Works of Aristotle. Revised Oxford Translation*. Princeton NJ: Princeton University Press.

Arnove, F. (1999) Introduction: Reframing Comparative Education: The Dialectic of the Global and the Local. In Arnove, F. & Torres, C. (eds) *Comparative Education: The Dialectic of the Global and the Local*, pp 1-25. Lanham: Rowman & Littlefield.

Ayer, A. J. (1964). *Man as a Subject for Science*. London: Athlone Press.

Barry, C. (1998) Qualitative Data Analysis Software: Atlas/ti and Nudist Compared. In *Sociological Research Online*, 3(3). Available on-line http://www.socresonline.org.uk/socresonline/3/3/4.html [2002, 26 September].

Beare, H. and Slaughter, R. (1995) *Education for the 21st Century*. London: Routledge.

Benn, S. I., & Peters, R. S. (1959). *Social Principles and the Democratic State*. London: Allen and Unwin.

Bogdan, R. & Knopp Biklen, S. (1992) Qualitative Research for Education: An Introduction to Theory & Methods (2nd ed). Boston: Allyn & Bacon.

Bosetti, L., Landry, D. & Miklos, E. (1989) Critical Perspectives on Educational Planning and Policy Analysis. In *The Canadian Administrator*, 29(2) November, pp. 1-7.

Bradford DeLong, J. (2001) The World's Income Distribution: Turning the Corner? Available on-line: http://econ161.berkeley.edu/TotW/world_income_dist.html [2002, Sept 6]

Bronowski, J. (1973). *The Ascent of Man*. London: British Broadcasting Corporation.

Burns, R. (1990) *Introduction to Research Methods*. Melbourne: Longman Cheshire

Burns, R. (1999) *Introduction to Research Methods (4th ed)*. Frenchs Forest: Longman

Carr, W. & Kemmis, S. (1986) *Becoming Critical: Knowing through Action Research*. Victoria: Deakin University Press.

Carspecken, P. (1996) Critical Ethrography in Education Research: A Theoretical and Practical Guide. New York: Routledge.

Connole, H. (2000) The Research Enterprise. In Smith B (ed) *Issues and Methods in Research (Study Guide) UED601/SG/01/ver2*. Underdale: The University of South Australia.

Couvalis, 1997, The Philosophy of Science. London: SAGE Publications.

Cribb, J. (2002) Share the Good Word on Unis. In *The Australian* (Higher Education supplement, Wednesday 5 June 2002).

de Landsheere, G. (1997) History of Educational Research. In J. Keeves (ed) *Educational Research, Methodology, and Measurement: An International Handbook (2nd ed)* Cambridge: Pergamon.

Doordan, A. (1998) *Research Survival Guide*. Philadelphia: Lippincott-Raven Publishers.

Ebison, M. G. (1993). Newtonian in Mind but Aristotleian at Heart. *Science & Education, 2*, 345-62.

Einstein, A., & Infeld, L. (1938). *The Evolution of Physics*. New York: Simon & Schuster.

Gibbons, J. A. (1979). Curriculum Integration. *Curriculum Inquiry, 9:4*, 322-32.

Grazino, A. & Raulin, M. (1996) *Research Methods: A Process of Enquiry (3rd ed)*. New York: Longman.

Husén, T. (1997) Research Paradigms in Education. In J. Keeves (ed) *Educational Research, Methodology, and Measurement: An International Handbook (2nd ed)* Cambridge: Pergamon.

Keeves, J. P. (1997). Models and Model Building. In Keeves, J. P. (Ed.), *Educational Research: Methodology and Measurement. An International Handbook*. (pp. 386-94). Oxford: Pergamon.

Keeves, J. (2002) Research in Education: Nature, Impact, Needs and Priorities. Unpublished draft.

Kukla, A. (2000). Social Constructivism and the Philosophy of Science. London: Routledge.

Lakomski, G. (1997) Critical Theory and Education. In Keeves, J. (ed) *Educational research, methodology, and measurement: an international handbook (2nd ed)*. Cambridge: Cambridge University Press.

Lombardi, O. (1999). Aristotleian Physics in the Context of Teaching Science: a Historical-Philosophical Approach. *Science & Education, 8*, 217-39.

McCalla-Chen, D. (1996) A Trifle Experience: Centre Ground Approach to Research. Draft paper for presentation at the BERA Conference, Lancaster University, 12-15 September 1996. Available on-line: http://www.leeds.ac.uk/educol/documents/000000124.doc [2002, 3 Sept].

MacIntyre, A. (1988) Whose Justice? Which Rationality? Indiana: University of Notre Dame Press.

Mirriam-Webster (2002) *Collegiate Dictionary*. Available on-line: http://www.m-w.com/cgi-bin/dictionary [2002, 12 September].

Paul, J. L., & Marfo, K. (2001). Preparation of Educational Researchers in Philosophical Foundations of Inquiry. *Review of Educational Research, Winter 2001 Vol 71, No. 4*, pp. 525-547.

QSR (Qualitative Solutions and Research) (2002) Products / Nvivo. Available on-line: http://www.qsr-software.com/products/nvivo.html [2002, 26 September].

Rizvi, F. (1986) Bureaucratic Rationality and the Possibility of Democratic Governance in Education. In *The Australian Administrator*, 7(1) February 1986, pp. 1-4.

Slade, M. (2002) Listening to the Boys: Issues and Problems Influencing School Achievement and Retention. In *Flinders University Institute of International Education Research Collection: Number 5*. South Australia: Shannon Research Press.

Smith B (ed) (2000) Issues and Methods in Research (Study Guide) UED601/SG/01/ver2. Underdale: The University of South Australia.

Smith, K. (1989) The Nature of Social & Educational Enquiry: Empiricism versus Interpretation. New Jersey: Ablex.

Stinner, A. (1994). The Story of Force: from Aristotle to Einstein. *Physics Education*, *29*, 77-85.

US Census Bureau (2002) *Total Midyear Population for the World: 1950-2050*. Available on-line: http://www.census.gov/ipc/www/worldpop.html [2002, 5 Sept].

Walker, J. C., & Evers, C. W. (1997). Research in Education: Epistemological Issues. In Keeves, J. P. (Ed.), *Educational Research, Methodology, and Measurement: an International Handbook 2nd ed*. (pp. 23-31). Oxford: Elsevier Science.

Welch F. (1999) The Triumph of Technocracy or the Collapse of Certainty? Modernity, Postmodernity, and Postcolonialism in Comparative Education. In Arnove, F. & Torres, C. (eds) *Comparative Education: The Dialectic of the Global and the Local*, pp. 25-49. Lanham: Rowman & Littlefield.

Williamson, K. (2000) *Research Methods for Students & Professionals: Information Management & Systems*. Wagga Wagga: Charles Sturt University Centre for Information Studies.

Stepping from single to mixed sex education: Boys' progress and perceptions during the restructuring

Shirley M. Yates
Flinders University School of Education *shirley.yates@flinders.edu.au*

In 1999 an independent school in South Australia took the courageous step of introducing coeducation into their single sex boys' school. Mixed sex education was phased in over two years, with girls admitted to Years 7 to 12 in 1999 and Years 3 to 6 in 2000. Progress and perceptions of school climate were measured in all primary and secondary level boys in the school in the inaugural year of coeducation and then annually over the next two years. Hierarchical Linear Modelling analyses revealed an interesting pattern of interrelationships between boys' progress and cohort and grade level group perceptions of satisfaction, cohesiveness and competition within the school. Perceptions of the difficulty of schoolwork were also significantly related to boys' progress during the restructuring period.

Boys, single-sex education, coeducation, progress, perceptions

INTRODUCTION

The issue as to whether it is more beneficial for boys to be educated together with girls or separately has been hotly debated in the research literature and in the press for some considerable time, with very little consensus being reached. The contention that coeducational learning environments are happier, friendlier, more pleasant and gregarious than single sex schools was first raised by Dale (1969, 1971, 1974) in his 26-year long study of grammar schools in England. He reported that the more positive, conducive climate in coeducational schools did not come at the expense of academic achievement. Claims of more positive climates in mixed sex education were substantiated by Schneider and Coutts (1982), but they did not concur with Dale's statements that academic achievement was unaffected. Many contradictory assertions have been made as to whether boys and girls are more successful academically in single sex institutions (Astin, 1977; Riordan, 1985; Lee & Bryk, 1986; Young and Fraser, 1990; Kelly 1996), coeducational institutions (Gilligan, 1982: Spender, 1982; Marsh, Smith, Marsh & Owens, 1988; Marsh, 1989), or single sex classes within coeducational schools (Rennie & Parker, 1997; Jackson & Smith, 2000). Other evidence has indicated that there are no differences in achievement in single and mixed sex settings (Miller & Dale, 1974; Rutter, Maughan, Mortimer & Outson, 1979) and that school type does not benefit either boys or girls once other variables are controlled (Marsh, 1989; Marsh & Rowe, 1996).

Student progress during the restructuring from single sex to mixed sex education has been examined in only a few studies, with the majority of these focussed at the secondary level. Jones and Thompson (1981) found no improvement in academic achievement when an all male private school became sex integrated. Similarly, there were no significant differences in achievement following the merger between a boys and girls single sex non-government secondary school (Marsh, 1989; Marsh, *et al*, 1988; Smith, 1994). By contrast, significant

increases in general educational progress were evident in primary and secondary boys from Grades 3 to 12 following the introduction of co-education into two non-government single sex boys schools (Yates, 2000; 2001a; 2001b; 2002a). Contradictory evidence has also been reported in relation to classroom conduct following the introduction of coeducation (Mael, 1998), with Jones and Thompson (1981) citing improvements and Payne and Newton (1990) mixed views about the behaviour of boys. In a separate study of a merger of a single sex secondary girls school with a boys school, girls reported feeling intimidated, hesitant, uncomfortable and dumb during mathematics lessons in mixed sex classrooms, but the views of the boys were not canvassed (Steinbeck & Gwizdala, 1995). Boys' educational progress in newly amalgamated mixed education classrooms has not been investigated in relation to perceptions of the school learning environment.

School climate has long been recognised as one of the variables having an important effect on student learning (Rutter, Maughan, Mortimer & Outson, 1979; Anderson, 1992; Fraser, 1994; Bulach & Malone, 1994: Bulach, Malone & Castleman, 1995). A large body of evidence attests to strong associations between cognitive and affective outcomes for students and their perceptions of the psychosocial characteristics of their classrooms (Haertel, Walberg & Haertel, 1981; Fraser, Welch, Hattie & Walberg, 1987; Fraser, 1998). A meta-analysis of studies involving 17,805 students in 823 classes in eight subject areas across four nations found student achievement was enhanced in classrooms with greater Cohesiveness, Goal Direction and Satisfaction and less Disorganisation and Friction (Haertel, Walberg & Haertel, 1981). Cohesiveness, Friction and Satisfaction were originally identified as significant aspects of Relationship dimensions and Goal Direction and Disorganisation as part of the System Maintenance and Change dimensions postulated by Moos (1974). In the same classification schema of human environments Moos included a third dimension of Personal Development encompassing Competitiveness and Difficulty. The Relationship, Personal Development and System Maintenance and Change dimensions have been studied in many different learning environments, predominantly at the secondary school level (Fraser, 1998), but have not been measured in schools during the introduction of mixed sex education into a single sex school.

The present study was conducted in a non-government fee-paying school during the period of restructuring from single to mixed sex education. Co-education was phased into the school over a two-year period, with 43 girls entering Grades 7 to 12 in Term 1, 1999 and 23 girls in the primary Grades 3 to 6 in Term 1, 2000. In total, across Grades 3 to 12 there were 92 girls in 2000 and 123 girls in 2001. Perceptions of the school learning environment and educational progress were measured annually in Term 4 in primary and secondary school boys during the two-year transitional period and in the following year. The study began with boys present at the time of the initial introduction of co-education, and followed them as they experienced the changes brought about by the radical alteration to the gender mix in the school population. Their perceptions, measured by questionnaires, tapped their impressions and experiences over several lessons, activities and events and when pooled across them, provided a rich source of direct information about life within the coeducational school.

AIMS

The aims of this study were to:

1. measure, monitor and evaluate boys' perceptions of school climate and educational progress during the restructuring from single to mixed sex education;
2. examine cohort and grade level differences in boys' perceptions of school climate and educational progress over time;

METHOD

Participants

All boys in Grades 3 to 11 attending the school in 1999, when coeducation was introduced into Grades 7 to 12, participated. Table 1 presents the numbers of boys in Grades 3 to 11 at Time 1 (T1) (1999), Grades 4 to 12 at Time 2 (T2) (2000) and Grades 5 to 12 at Time 3 (T3) (2002).

Table1. Number of boys by Grade level at Time 1, Time 2 and Time 3

	Gr 3	Gr 4	Gr 5	Gr 6	Gr 7	Gr 8	Gr 9	G 10	G11	G 12	Total
T1	25	31	34	40	40	49	71	75	76		**441**
T2		23	27	33	40	40	48	71	76	76	**433**
T3			22	23	27	29	32	43	48	57	**363**
Total	**25**	**54**	**83**	**98**	**107**	**118**	**151**	**189**	**199**	**133**	**1237**

There is some variability in the numbers of boys over the three-year period, with higher levels of attrition evident in some grades than in others. Student ID numbers used for data entry in SPSS (Norusis, 1993) were generated from school class lists at T1. Some boys were absent at T1, T2 or T3 or had left the school altogether at T2 or T3. There were, however, sufficient numbers of boys at each grade level on each occasion for the statistical analyses to be conducted.

Instruments

School Climate Questionnaires:

Boys' perceptions of the psychosocial climate of the school were measured with *My School Inventory*, adapted from *My Class Inventory* (see, Fisher & Fraser, 1981; Fraser, Anderson & Walberg, 1982) or the *School Learning Environment Inventory* a modification of the *Learning Environment Inventory* (Anderson & Walberg, 1974; Fraser, Anderson & Walberg, 1982). In each inventory boys were presented with a series of statements and were asked to indicate whether or not they agreed with them. *My School Inventory* was composed of 38 items measuring five aspects of school climate while the *School Learning Environment* contained 108 items measuring 15 aspects. Five aspects, defined as separate subscales, measuring boys' perceptions of the Relationship and Personal Development dimensions (Moos, 1974) of the school were common to both inventories. The Relationships dimensions, measured through three subscales of Cohesiveness, Friction and Satisfaction in each inventory, tapped the nature and intensity of students' personal relationships, conflict, arguments and disagreements between students and their contentment and happiness with the school learning environment respectively. The Personal Development dimensions measured by two subscales of Competitiveness and Difficulty assessed the extent to which students competed with each other and their perceptions of the difficulty of their schoolwork. Sample items from the Relationship and Personal Development dimensions within *My School Inventory* (MSI) and *School Learning Environment Inventory* (SLEI) are presented in Table 2.

Educational Progress:

Students' educational progress was measured over three consecutive years with the *Word Knowledge Test 1, 2* or *3* (Thorndike, 1973). These tests have been used previously in international studies as a general assessment of students' achievement and underlying scholastic aptitude. They were chosen as they could be administered, compiled into a single

scale with the common items linking procedure and compared across Grades 3 to 12. Each test consists of 40 word pairs, rated as the same or opposite in meaning. Thirteen word pairs are common to Tests 1 and 2 and 20 items are common to Tests 2 and 3. One word pair is common to all three tests. Table 3 presents sample items from Test 1, Test 2 and Test 3.

Table 2. Sample items for the Relationship and Personal Development dimensions within *My School Inventory* (MSI) and *School Learning Environment Inventory* (SLEI)

Relationship Dimensions:		
Cohesiveness		
MSI	**Item 28**	***Children in our school like each other as friends***
SLEI	**Item 31**	***Students in the school are frequently personal friends***
Friction		
MSI	**Item 2**	***Children are always fighting with each other***
SLEI	**Item 6**	***There is constant bickering among students***
Satisfaction		
MSI	**Item 16**	***Most children say the school is fun***
SLEI	**Item 73**	***Students look forward to coming to classes***
Personal Development Dimensions:		
Competitiveness		
MSI	**Item 24**	***Some students try to do their work better than the others***
SLEI	**Item 30**	***Students compete to see who can do the best work***
Difficulty		
MSI	**Item 3**	***In our school work is hard to do***
SLEI	**Item 54**	***Students tend to find the schoolwork hard to do.***

Table 3. Sample items from Word Knowledge Test 1, Test 2 and Test 3

Item	***Word Knowledge Test 1***		***Word Knowledge Test 2***		***Word Knowledge Test 3***	
1	quick	slow	savoury	insipid	acquire	dispel
40	scarce	rare	obvious	indisputable	ephemeral	eternal

Procedure

The school climate questionnaire and word knowledge test were administered to all boys in their classrooms at the same time and day in October at T1, T2 and T3. *Word Knowledge Test 1* was administered to boys in Grades 3 to 7, *Word Knowledge Test 2* to Grades 8 to 10 and *Word Knowledge Test 3* to Grades 11 and 12. Primary school boys in Grades 3 to 7 were administered *My School Inventory*, with the *School Learning Environment Inventory* given to the secondary level boys in Grades 8 to 12. *My School Inventory* was also completed by Grade 8 and Grade 9 boys to provide a common group for equating purposes.

Analyses

Boys' responses at T1, T2 and T3 were entered into an SPSS file (Norusis, 1993), with data matched across each occasion through their student ID number. The two school climate questionnaires and three word knowledge tests were calibrated with the Rasch scaling procedure (Rasch, 1966) using QUEST (Adams & Khoo, 1994) and all non-fitting items deleted. Five separate school climate subscales of Cohesiveness (COH), Competitiveness (COM), Difficulty (DIF), Friction (FRI) and Satisfaction (SAT) were then formed from the designated items in the *My School Inventory* and *School Learning Environment Inventory.*

Each of the subscales from the two inventories were linked by the responses of the Grade 8 and 9 boys who had completed both questionnaires. The combined COH subscale contained 13 items, the COM subscale 12 items, the DIF subscale 14 items, the FRI subscale 15 items and SAT subscale 14 items.

A single *Word Knowledge* (WK) scale of student progress was formed from *Word Knowledge Tests 1, 2 and 3,* with the tests linked by the 13 items common to Tests 1 and 2, the 20 items common to Test 2 and 3 and the one item common to all three tests. The scoring of WK was then anchored to only those students who answered all items. Case estimate scores for all boys for T1, T2 and T3 were equated concurrently for the COH, COM, DIF, FRI and SAT scales and the WK scale.

Relationships between the five school climate subscales and the WK scale measured over the three years of the study were analysed with Hierarchical Linear Modelling (HLM) (Raudenbush, Bryk, & Congdon, 2000), using HLM5. This procedure is eminently suited to longitudinal designs (Von Eye, 2001). Within group comparisons were made over time at level-1, and between cohort groups and grades at level-2. Cohort groups consist of the same boys clustered by their initial grade at T1, while Grade groupings are composed of boys in that grade level at T1, T2 and T3 respectively.

RESULTS

Students in any school or classroom are more similar to each other than they are to students randomly sampled from a school region, sector or national population (Osborne, 2000). Furthermore, as this study was conducted in a single fee paying, non-government school, it is likely that students share certain background, environmental and experiential characteristics (Osborne, 2000). Hierarchical Linear Modelling was therefore employed as it permits the examination of the direct effect of various potential predictors at level-1 and level-2 as well as modelling of cross-level interaction effects. Two models were developed, with boys grouped at level-2 by cohort in Model 1, presented in Figure 1 and by grade level at T1, T2 and T3 in Model 2, presented in Figure 2. Boys' progress, measured by WK, was designated as the outcome variable in both models. Coefficients and standard errors for each significant variable, which is enclosed within an ellipse, are presented at both the student and group level within each model. Taken together, all of the five school climate subscales are significantly related to boys' progress at either level 1 or level 2 or both, but the manner in which they do so varies between the two models.

Model 1:Boys grouped by cohort

Figure 1 and Table 4 present the significant effects for Model 1.

At level 1 boys' perceptions of the difficulty of their school work (DIF) and Time are significantly related to WK. On average, all cohort groups show significant increases in WK scores over time. At level 2 there is a direct, significant, negative relationship between Cohort 9 and WK, indicating that this cohort group had significantly lower WK scores than the other cohort groups. There are also direct, significant, negative relationships between Satisfaction (SAT) and WK and between Competitiveness (COM) and WK. Cohort groups that on average had higher levels of SAT recorded lower WK levels, as did cohort groups with higher levels of COM. Conversely, cohorts which expressed lower levels of SAT experienced higher gains in WK. although the significant positive interaction effect between SAT and Time in Model 1 indicates that satisfaction increased across the school over the three years of the study. There is a significant positive interaction effect between SAT at level 2 and COH at level 1 indicating that boys reporting high levels of cohesiveness in cohorts that are, on average, more satisfied with school life have lower WK scores. By

contrast, less cohesive boys in cohort groups that express lower levels of satisfaction experience higher WK gains. However, cohesiveness alone is not a significant variable except for boys in Cohort 7 where it is negatively related to WK. Perceptions of difficulty of schoolwork (DIF) is a significant variable at both levels 1 and 2. The interaction effect between DIF at level 2 and level 1 indicates that the cohort group's perceptions of difficulty influence individual perceptions, such that on average, boys in cohort groups who do not perceive school work to be difficult are more likely to have higher WK scores. On the other hand, in cohort groups that perceive the work to be difficult there is almost no difference in the word knowledge scores of individual boys irrespective of their own perceptions of difficulty.

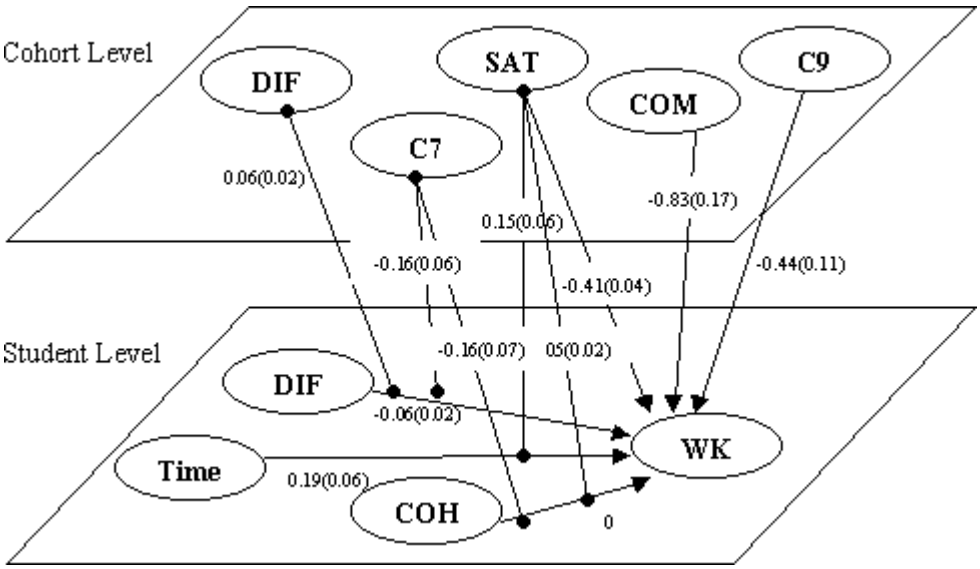


Figure 1. Model 1: Boys grouped by cohort

Table 4. Model 1: Boys grouped by cohort

Final estimation of fixed effects:					
Fixed Effect	**Coefficient**	**Std Error**	**T-ratio**	**d.f.**	**P-value**
For INTRCPT1, B0					
INTRCPT2, G00	1.18	0.04	30.28	5	0.000
ENRSSS_1G01	-0.41	0.04	-10.63	5	0.000
ENRSSC_1G02	-0.83	0.17	-4.84	5	0.003
C9_1, G03	-0.44	0.11	-4.02	5	0.014
For TIME slope, B1					
INTRCPT2, G10	0.19	0.06	3.10	7	0.019
ENRSSS_1,G11	0.15	0.06	2.68	7	0.032
For ENRSSDIF slope, B2					
INTRCPT2, G20	-0.06	0.02	-2.59	6	0.041
ENRSSD_1, G21	0.07	0.03	2.30	6	0.061
C7_1, G22	-0.16	0.06	-2.60	6	0.041
For ENRSSCOH slope, B3					
INTRCPT2, G30	0.00	0.02	-0.09	6	0.931
ENRSSS_1,G31	-0.05	0.02	-2.42	6	0.051
C7_, G32	-0.16	0.07	-2.26	6	0.064
Final estimation of variance components:					
Random Effect	**Standard Deviation**	**Variance Component**	**df**	**Chi-square**	**P-value**
INTRCPT1, U0	0.07	0.00	5	7.66	0.175
TIME slope, U1	0.15	0.02	7	19.10	0.008
ENRSSDIF slope, U2	0.01	0.00	6	2.29	>0.500
ENRSSCOH slope, U3	0.04	0.00	6	9.70	0.137
level-1, R	0.86	0.75			

Model 2:Boys grouped by Grade level

Figure 2 and Table 5 present the significant effects for Model 2.

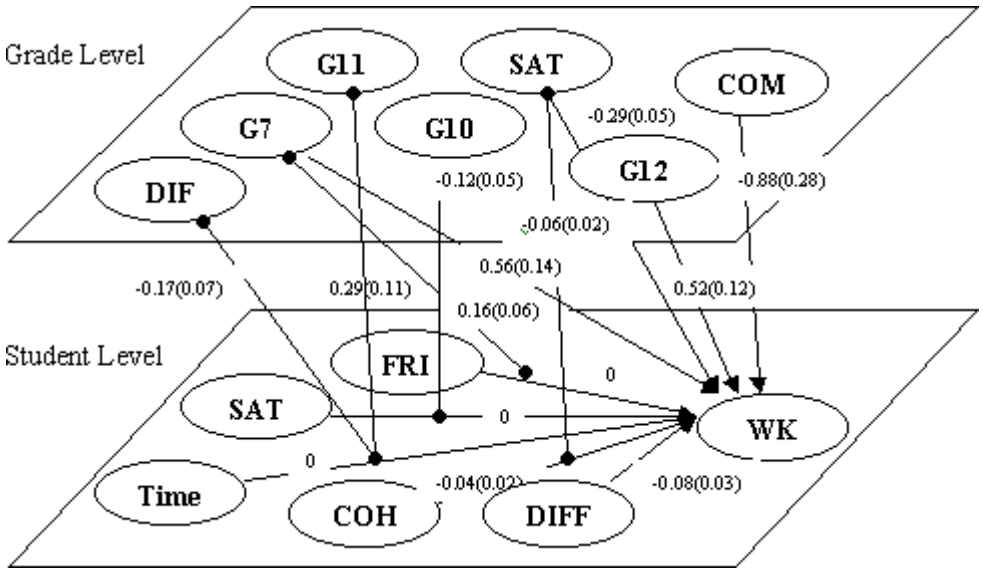


Figure 2. Model 2: Boys grouped by grade

Table 5.Model 2: Boys grouped by grade

Final estimation of fixed effects					
Fixed Effect	**Coefficient**	**Std. Error**	**T-ratio**	**d.f.**	**P-value**
For INTRCPT1, B0					
INTRCPT2, G00	0.93	0.06	16.22	5	0.000
ENRSSS_1G01	-0.29	0.05	-6.27	5	0.000
ENRSSC_1G02	-0.88	0.28	-3.11	5	0.030
Y7_1, G03	0.56	0.14	4.15	5	0.012
Y12_1. G04	0.52	0.12	4.13	5	0.012
For TIME slope, B1					
INTRCPT2, G10	0.01	0.05	0.13	7	0.898
ENRSSD_1, G11	-0.17	0.07	-2.38	7	0.049
Y11_1, G12	0.29	0.11	2.54	7	0.039
For ENRSSSAT slope, B2					
INTRCPT2, G20	0.03	0.02	1.25	8	0.247
Y10_1, G21	-0.12	0.05	-2.19	8	0.059
For ENRSSFRI slope, B3					
INTRCPT2, G30	0.000	0.02	0.22	8	0.834
Y7_1, G31	0.16	0.06	2.70	8	0.027
For ENRSSDIF slope, B4					
INTRCPT2, G40	-0.08	0.03	-2.63	9	0.028
For ENRSSCOH slope, B5					
INTRCPT2, G50	-0.04	0.02	-2.02	8	0.078
ENRSSS_1, G51	-0.06	0.02	-3.67	8	0.007

Final estimation of variance components:					
Random Effect	**Standard Deviation**	**Variance Component**	**d.f**	**Chi-square**	**P-value**
INTRCPT1, U0	0.12	0.02	4	19.72	0.001
TIME slope, U1	0.07	0.01	6	7.22	0.300
ENRSSSAT slope, U2	0.02	0.00	7	2.77	>.500
ENRSSFRI slope, U3	0.03	0.00	7	6.72	>.500
ENRSSDIF slope, U4	0.07	0.00	8	16.45	0.036
ENRSSCOH slope, U5	0.01	0.00	7	2.25	>.500
level-1, R	0.86	0.74			

All five of the school climate variables are evident at either level 1 or level 2 in Model 2. In these analyses in which boys were grouped by grade, WK scores for all grade levels were not significantly different from the overall mean, except for Grades 7 and 12 which were significantly higher. With the exception of Grade 11, there were no significant increases in WK scores over time. Average WK scores for each grade were influenced by average levels of SAT and COM, so that grades with higher scores in SAT and COM respectively tended to have lower WK scores. An interaction was also evident between SAT and COH. Boys in grades which, on average, recorded higher levels of SAT and COH had lower WK scores. By contrast, those in Grades which on average had lower levels of SAT and COH gained higher WK scores. At the individual student level, satisfaction with school life was not a significant factor except for students in Grade 10 where higher levels of SAT were related to less progress. Boys' perceptions of friction within the school was similarly not a significant factor in the grade analyses except in Grade 7 where FRI had a positive effect on WK. However, there was a significant interaction between DIF and Time, such that grade levels which on average perceived the perceived schoolwork to be less difficult had increasing WK scores over time. Conversely, grade levels where on average schoolwork was perceived to be more difficult evidenced less progress over time. As boys' WK scores decreased, they perceived their schoolwork to be more difficult.

Summary of the results

1. All cohort groups made significant progress as measured by WK. Grades 7 and 12 had significantly higher WK scores when compared to the overall mean, but only Grade 11 recorded significant increases over time.
2. Boys' WK scores were influenced by their perceptions of the school climate at both the individual and group level.
3. Satisfaction and competitiveness were significant variables in both the cohort and grade analyses.
4. Satisfaction also interacted with cohesiveness to influence WK outcomes.
5. Perceptions of the difficulty of schoolwork had a significant influence on boys' progress, particularly for cohort 7 and over time in the grade analyses.
6. Friction was not a significant variable except for Grade 7.

DISCUSSION

In a comprehensive review of the relationships between socioemotional and academic development in single and mixed sex education, Mael (1998) foreshadowed the need for future research to take both the individual student and organisational contexts of an educational setting factors into account. While the present study was conducted only on a single school during a period of significant structural change, it was designed so that individual and group factors could be considered together with aspects of the learning environment. The use of Hierarchical Linear Modelling enabled predictive relationships between boys' perceptions of the school climate and their educational progress to be explored at both individual and group levels, as well as any potential cross level interactions between them. Furthermore, while previous studies have been focussed predominantly on students at the secondary level, both primary and secondary boys participated in this study. With the use of the Rasch scaling procedure, single scales were formed which enabled boys' perceptions and progress to be viewed from Grades 3 to 12 across the school as well as over time. Finally, this study is unique precisely because it measured boys' progress and captured

their perceptions of the climate of their school during the period of the restructuring from a unisex to a coeducational school.

The finding that educational progress, as measured by the Word Knowledge test, increased significantly for all cohorts indicates that over time, the gradual introduction of girls into a single sex boys school was not disadvantageous for boys present in the school when coeducation was introduced. However, this progress was clearly influenced by boys' perceptions of climatic variables - most notable, satisfaction, cohesion, competitiveness and difficulty. The relationships between satisfaction and progress in Models 1 and 2 affirm and extend the findings from the meta-analytic review of Haertel, Walberg & Haertel (1981). The fact that the relationship between the two variables in both models is negative can be interpreted in light of previous analyses of the longitudinal data from this study which has shown boys' satisfaction with school life to decrease across Grade levels and word knowledge scores to increase (Yates, 2001b, 2002b). In general, boys in the higher grade levels are less satisfied with school than boys in the lower grades, although the cohort analyses in Model 1 show satisfaction to have increased significantly over time. However, there are also interesting interaction effects between boys' progress and Moos' (1974) Relationship dimensions of satisfaction and cohesion in both models. By itself cohesion was not a significant variable except for boys in Cohort 7. Likewise, with the exception of Grade 7, the Relationship dimension of friction was not a significant variable. These results are somewhat surprising, as the meta-analysis of classroom climate research highlighted the important roles played by Cohesiveness, Friction and Satisfaction in student achievement outcomes (Haertel, Walberg & Haertel, 1981). While this study was confined to boys only, these results would suggest that the manner in which Moos (1974) Relationship dimensions operate may need to be reconsidered, particularly when group effects are taken into account with newer statistical techniques such as Hierarchical Linear Modelling (Mael, 1998).

In both models there are significant, negative relationships between boys' perceptions of competitiveness in relation to their schoolwork at both cohort and grade levels and their educational progress at level 1. Similarly, there are signification interactions between boys' perceptions of the difficulty of schoolwork and progress over time. Both of these variables belong to Moos' (1974) dimension of Personal Development. These findings need to be investigated further to determine whether they are gender specific, or whether they pertain only to the present sample of boys who were attending the school at the time that coeducation was first introduced. Explication of this result may have implications for the widely held view, consistent across several studies, that boys are generally advantaged in coeducational settings (see, Mael, 1998).

Overall, the results from this study found boys' educational progress and perceptions of their school climate during the transition from single sex to coeducation to be interrelated in a manner not previously demonstrated in earlier research. The HLM analyses showed that these relationships operated at both the individual and group levels, with some differential effects for cohort and grade level groupings. These findings contribute to the debate about single sex and coeducation, particularly in relation to the transition from one system to another, as they indicate boys' progress is influenced significantly by some school climate variables. They also establish the importance of examining factors at both the individual student and group levels. Further studies are now needed to confirm whether these interrelationships are sustainable over time, particularly when the initial impact of the restructuring from single to mixed sex education has attenuated in the school. The study could also be extended to include other variables, the girls who have entered the school, as well as the perceptions of the teachers.

Acknowledgments

1. Professor John P. Keeves for his unfailing interest, support and expertise.
2. I Gusti Ngurah Darmawan for invaluable assistance with data analyses.
3. Research assistants John Harris & Dianne Harris.
4. The research was supported by Australian Research Council and Flinders Small Grants.

REFERENCES

Adams, R.J. & Khoo, S.K. (1994). *Quest- The interactive test analysis system.* Hawthorn, Victoria: ACER.

Anderson, C. S. (1992). The search for school climate: A review of research. *Review of Educational Research, 32,* 368-420.

Anderson, G. J. & Walberg, H. J. (1974). Learning environments. In H. J. Walberg (Ed.), *Evaluating educational performance: A sourcebook of methods, instruments, and examples.* Berkeley: McCutchan.

Astin, A .W. (1977). On the failure of educational policy. *Change, 9,* 40-45.

Bulach, C. R. & Malone, B. (1994). The relationship of school climate to the implementation of school reform. *ERS Spectrum, 12* (4), 3-8.

Bulach, C. R., Malone, B. & Castleman, C. (1995). An investigation of variables related to student achievement. *Mid-Western Educational Researcher 8* (2), 23-29.

Dale, R. R. (1969). *Mixed or single sex schools?: Vol 1 A research study about pupil-teacher relationships.* London: Routledge& Kegan Paul.

Dale, R. R. (1971) *Mixed or single sex schools? Vol 2 Some social aspects.* London: Routledge & Kegan Paul.

Dale, R. R. (1974) *Mixed or single sex schools? Vol 3 Attainments, attitudes and overview.* London: Routledge & Kegan Paul.

Fisher, D. L. & Fraser, B. J. (1981). Validity and use of My Class Inventory. *Science Education, 65,* 145-156.

Fraser, B. J. (1994). Research on classroom and school climate. In D, Gabel (Ed.), *Handbook of research on science teaching and learning.* (pp. 493-541) New York: Macmillan.

Fraser, B. J. (1998). Science learning environments: Assessment, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *The international handbook of science education* (pp. 527-564). Dordrecht, The Netherlands: Kluwer.

Fraser, B. J., Anderson, G. J. & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and MY Class Inventory (MCI)* (third version) Perth: Western Australian Institute of Technology.

Fraser, B. J., Welch, W. W., Hattie, J. A. & Walberg, H. J. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11,* 147-252.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Havard University Press.

Haertel, G. D., Walberg, H. J. & Heartel, E. H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7,* 27 -36.

Jackson, C. & Smith, I. D. (2000). Poles apart? An exploration of single-sex and mixed sex educational environments in Australia and England. *Educational Studies, 26* (4), 409-422.

Jones, M. B. & Thompson, D. G. (1981). Classroom misconduct and integration by sex. *Journal of Child Psychology and Psychiatry. 22,* 401-409.

Kelly, A. (1996). Comparing like with like *Education, 187,* 14-15.

Lee, V. E. & Bryk, A. S. (1986). Effects of single-sex secondary schools on student achievement and attitudes, *Journal of Educational Psychology, 78,* 381-395.

Mael, F. A. (1998). Single-sex and coeducational schooling: Relationships to socioemotional and academic development. *Review of Educational Research, 68* (2), 101-129

Marsh, H. W. (1989). Effects of attending single-sex and coeducational high schools on achievement, attitudes, behaviours and sex differences., *Journal of Educational Psychology, 81,* 70-85.

Marsh, H. W. & Rowe, K. (1996). The effects of single-sex and mixed-sex mathematics classes within a coeducational school: A re-analysis. *Australian Journal of Education, 40,* 147-162.

Marsh, H. W., Smith I. D., Marsh, M., & Owens, L. (1988). The transition from single-sex to coeducational high schools: Effects on multiple dimensions of self-concept and on academic achievement. *American Educational Research Journal, 25,* 237-269.

Miller, P, McC. & Dale, R. R. (1974). A comparison of the degree results of university students from coeducational and single-sex schools. *British Journal of Educational Psychology, 44,* 307-308.

Moos, R. H. (1974). *The social climate scales: An overview*. Palo Alto, California: Consulting Psychologists Press.

Norusis, M. J. / SPSS Inc. (1993). *SPSS for windows base system user's guide release 6.0.*

Osborne, J. W. (2000). Advantages of hierarchical linear modelling. *Practical Assessment, Research & Evaluation, 7* (1), [Available on-line: http://ericae.net/pare/getvn.asp?v=7&n=1]

Payne, M. A. & Newton, E. H. (1990). Teachers and students perceptions of major advantages and disadvantages of coeducational secondary schooling. *Australian Journal of Education, 34,* 67-86.

Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology, 19,* 49-57.

Raudenbush, S., Bryk, A. & Congdon, R. (2000). *HLM5 Hierarchical linear and nonlinear modelling.* Scientific Software International.

Rennie, L. J. & Parker, L. H. (1997). Perceptions of single- sex and mixed-sex mathematics classes. *Mathematics Education Research Journal. 9,* (3), 257-273.

Riordan, C. (1985). Public and catholic schooling: The effects of gender context policy. *American Journal of Education*, **5**, 518-540.

Rutter, M., Maughan, B., Mortimore, P. & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*, London: Open Books.

Schneider, F. W. and Coutts, L. M. (1982). The high school environment: A comparison of coeducational and single sex schools. *Journal of Educational Psychology, 74*, 898-906.

Smith, I. D. (1994). The coeducational/ single sex schooling debate. *Forum of Education, 51*, 15-31.

Spender, D. (1982). *Invisible women: The schooling scandal.* London: Writers and Readers Press.

Steinbeck, M. and Gwizdala, J. (1995) Gender differences in mathematics attitudes of secondary school students. *School Science and Mathematics. 95* (1), 36-42.

Thorndike, R. L. (1973). *Reading comprehension in fifteen countries.* New York, Wiley: and Stockholm: Almqvist & Wiksell.

Von Eye, M. (2001). *Structural Equation Modelling*. Workshop presented at the 9th European Association for Research in Learning and Instruction, Fribourg, Switzerland, 27 August. *Journal of Education, 43* (2), 142-156

Yates, S. M. (2000). *Student optimism and pessimism during the transition to co-education*. Presented at the annual conference of the Australian Association for Research in Education, Sydney, December, www.aare.edu.au/00pap/yat00439.htm.

Yates, S. M. (2001a) *Students' educational achievement and perceptions during the transition from single sex to co-education*. Paper presented at the 9th European Association for Research in Learning and Instruction, Fribourg, Switzerland, 29 August.

Yates, S. M. (2001b). Students' achievement and perceptions of school climate during the transition from single sex to coeducation. *International Education Journal, 2,* (4), *Educational Research Conference 2001 Special Issue*

Yates, S. M. (2002a). *Boys' academic progress during the reform of two single sex schools into co-educational schools.* Paper presented at the Second International Forum on Education Reform: Key Factors in Effective Implementation, Bangkok, 4 September.

Yates, S. M. (2002b). *Pulteney Grammar School. Report for Data Collected October 1999, 2000, 2001*. Flinders University, August 2002

Young, D. J. & Fraser, B. J. (1990). Science achievement of girls in single-sex and coeducational schools. *Research in Science and Technological Education, 8*, 5-20.

Overcoming Transactional Distance as a Barrier to Effective Communication over the Internet

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

The use of the Internet as a medium for distance education requires students to transmit large amounts of socio-emotional information to people who are often total strangers. This can make people feel vulnerable, so they limit the amount of this type of information they transmit across the Net, which in turn creates a barrier to effective communication. This study assessed the quality of the communication that occurred as part of the teaching in a subject in a Doctorate of Education program. The quality of the communication did not accord with the expectations gleaned from the literature, due to the nature of the relationships that existed between the participants. The findings of this study can be used as the basis of interventions designed to improve the effectiveness of the Internet as a medium for undergraduate distance education.

Distance education, transactional distance, discourse communities

INTRODUCTION

Verbal communication is supported by a raft of non-verbal signs and cues that reinforce what we are saying or clear up any ambiguities. For example, we may cross our arms when we feel threatened by what somebody else is saying, or we nod our heads when we agree with what they are saying. However, barriers to communication may arise because people from different cultural groups sometimes have different forms of non-verbal communication. The confusion that can be caused by seeing non-verbal communication that does not accord with what you are saying can lead to misinterpretation. Consequently, when we talk to other people we also need to engage with them non-verbally in order to reduce the transactional distance that exists between people.

Most users of email are probably aware of the problems that can be caused by transactional distance. When we use the Internet or email as a medium of communication we cannot use non-verbal forms of communication to help us get our messages across to our readers, sometimes with dire consequences. Scott Fahlman of the Department of Computer Science at Carnegie Mellon University was the first person to successfully address this problem. Fahlman is reported as saying that he saw the chaos that resulted from misunderstood postings. He is quoted as saying "the problem was that if someone made a sarcastic remark, a few readers would fail to get the joke, and each of them would post a lengthy diatribe in response. That would stir up more people with more responses, and soon the original thread of the discussion was buried" (Gengler 2002; 3). So on 19 September 1982 Fahlman sent the world's first emoticon, :-), the smiley face with which we are all familiar. The second emoticon was the grumpy face, :-(, followed soon afterwards. Fahlman intended that these two emoticons be inserted before a statement to indicate that the statement as intended to be taken light heartedly or seriously. However, the grumpy face soon evolved into a marker for displeasure, frustration or even anger. Emoticons are comprised of ordinary keyboard characters that form an image when viewed sideways. Within a few months lists of dozens of emoticons began to appear in emails and on Internet sites. Emoticons therefore, substitute

for non-verbal forms of communication, which facilitates communication across the Internet.

Transactional distance is becoming increasingly recognised as a barrier to the effective use of the Internet as a mode of distance education delivery. The highly personal nature of the communication that occurs when subjects are delivered over the Internet means that students are required to communicate significant amounts of socio-emotional information. The need to transmit large quantities of socio-emotional information to strangers over an impersonal mode of communication can make people feel vulnerable and open to personal attacks. So they limit the amount of this type of information that they include in their email messages or bulletin board postings, which in turn creates a barrier to communication. Emoticons are useful substitutes for non-verbal communication in emails, but is their use sufficient for Internet moderated teaching activities? This study investigated the variety of strategies that a class of postgraduate students used to reduce the transactional distance when using the Internet as a substitute for face-to-face seminars.

This paper has three further sections. The first section of this paper is a brief discussion of the problems associated with transactional distance in the context of distance education. The second section is a report of the cyber-colloquium, which was a study of the methods that doctoral students used to maximise the effectiveness of the Internet as a substitute for face-to-face teaching activities. The third section is a brief checklist for action when using the Internet as a teaching tool. This section incorporates some of the lessons that were learned during the cyber-colloquium.

TRANSACTIONAL DISTANCE

The University of South Australia, like many other Australian universities, is a second-generation distance education institution. These institutions are characterised by their reliance on the use of printed materials as the primary teaching medium (Jones 1996). As opposed to third generation distance education institutions who use the Internet as their primary teaching medium. The major objective of second generation distance education organisations is the production and distribution of print based teaching materials. Consequently, little attention is paid to the concept that learning is a social process that relies heavily on communication between students and between students and their teachers.

Print-based media rely heavily on the postal system to distribute teaching materials and return of student assignments. However, the increasingly international nature of the distance education services that are offered by Australian universities, such as the University of South Australia, and the continued reliance on postal services results in significantly increased time to distribute teaching materials. These increased times can lead to significant learning problems for distance students. Jones (1996) identified the following eight problems;

- the late arrival of teaching materials;
- inability to respond to errors or omissions in the teaching materials;
- inability to respond timely to student concerns;
- delays in providing both staff and students with feedback on the progress of students;
- a general lack of communication between students and their peers and teachers which results in feelings of isolation;
- long lead times in developing teaching materials;
- little control by teaching staff over the style and presentation of teaching materials; and
- an inability to cope with change.

Despite these obvious problems print based media remains popular with Australian distance education providers for a variety of reasons.

Distance educational students are commonly disadvantaged by both geographical and transactional distance (Caladine 1993). Transactional distance is the psychological space that exists between the learner and the teacher. It is a function of the extent of the dialogue that students have with other students and their teachers as well as the responsiveness of the subject to respond to the needs of individual students. The print-based nature of second-generation distance education institutions is not only static, but it also limits the dialogue between students and their teachers. It also reduces the responsiveness of the subject. These problems with print based teaching materials results in large transaction distance for distance education students.

Harasim *et al* (1995) argue that teaching across networks, which of course includes the Internet, allows students to communicate with their teachers and other students outside of normal class times or business hours in order to;

- extend opportunities for class discussion and debate;
- increase access to teachers;
- submit or exchange assignments;
- facilitate group tasks; and
- expand opportunities for informal group discussion and social interaction.

In the context of distance education, networking and the use of Internet tools such as WebCT can be used to enhance;

- communication between students and teachers;
- information exchange;
- assignment submission; and
- feedback.

Networking also introduces opportunities for improved interaction between students and collaborative learning opportunities in distance education.

Consequently, both geographical and transactional distance can be overcome by using the Internet as a medium for distributing teaching materials and as a forum for learning activities. The Internet also permits distance education students to engage in the informal learning activities that on-campus students take for granted. However, the increased use of the Internet is unlikely to be extended to point where it supplants the use of printed materials in second-generation distance education institutions due to the cost of hardware and the cost of providing additional information technology support personnel. For example, UniSANET, the Internet platform that was recently introduced at the University of South Australia does not include a function like WebCT due to the cost of installing sufficiently powerful servers. Hence, the coordinators of subjects that are taught externally cannot conduct bulletin boards or chat sessions.

However, it is possible that the geographical and transactional distance between staff and students might be reduced through better use of email. For example, at the University of South Australia it is becoming increasingly common for subject coordinators to set up web serve lists of student email addresses. These lists serve a variety of functions, but they are primarily used as a medium by which information and teaching resources are transmitted to students by the subject coordinator. However, they can also be used as a mode of communication between staff and the class as a whole. This ability of web serve lists to facilitate communication between staff and students may act to reduce the transactional distance between staff and students who are studying in both the internal and external mode,

but not to the extent that web based communication tools allow. Consequently, web serve lists might be a low cost, *albeit* an inferior, alternative.

THE CYBER COLLOQUIUM

Introduction

The participants of this study were the dozen students enrolled in *Doctoral Colloquium*, which is a compulsory subject in the Doctorate of Education at the Flinders University of South Australia, and the subject coordinator. This is a full year subject that provides final year doctoral students with an opportunity to reflect on the research projects that they had undertaken during their doctoral studies. The primary teaching activity is a two-hour weekly seminar in which students present their own work for discussion. The cyber-colloquium was designed to replace the last seminar for the subject. The participants logged onto WebCT from their homes and offices, while the session was facilitated by the subject coordinator from his office at the university. The cyberspace colloquium was conducted in two phases. The first was a bulletin board that was open from Wednesday 3 November to Sunday 7 November 1999. During that period, students were required to post at least one comment to the bulletin board each day. The second phase was a one-hour synchronous chat session held on Monday 8 November in place of the usual seminar for this subject. The class was given a two-hour training session on the use of WebCT in the week prior to the chat session being conducted. Participants were asked to read the report entitled *Community Partnerships in Education 1998, Volume 1* (DETE 1998), which is commonly referred to as *Partnerships 21*, from the perspective of a member of a school council who has been asked to report back to his or her school council to recommend whether the school should accept the State Government's invitation to participate in the *Partnership 21* initiative. The bulletin board and chat session were designed as fora in which the students in this subject could discuss the issues, develop their arguments and arrive at their conclusions.

The focus of the study was not a report on the merits, or otherwise, of *Partnerships 21*. Rather, the focus of this study is to report on the quality of the communication that took place during both phases of the cyber colloquium and to make recommendations about how the quality of the communication could be improved.

The Bulletin Board

The comments posted to the bulletin board were, by and large, quite long and quite formal. These comments tended to be at least one, but often two or more long paragraphs in length. The paragraphs resembled the building blocks of good essays. They tended to be well constructed with lead sentences, a formal argument and a concluding sentence. Indeed, a report about the merits or other wise of *Partnerships 21* could easily be written simply by collating these paragraphs into a single document.

The report that the class was asked to read clearly polarised the group. Most did not favour 'opting in'. However, at least two members of the group did favour 'opting in'. From reading the comments it became clear that that contributions to the discussion reflected the value systems and political beliefs of the contributors. The group is too highly educated for the comments to be simple assertions of their point of view on the discussion topic. Consequently, these paragraphs tended to be argumentative and supportive of the writer's own views. Hence they often contained references to texts in order to provide evidence to support the argument of the writer.

The participants were using time and space very differently during the bulletin board phase than they did during the face-to-face seminars. The physical separation of the class members

and the asynchronous nature of bulletin boards provided participants with the time required to carefully craft a well constructed comment or response that reflected their points of view. The bulletin board format also meant that contributors to the discussion could not be interrupted. However, it needs to be noted that the participants in this study were all mid-career educators with good written communication skills who were employed in one bureaucracy or another. Hence, they are highly experienced at communicating complex ideas through the medium of short pieces of writing, such as letters and memoranda. At this stage it is important to note that the written communication skills of the participants to this study are, hopefully, much better than then those of the vast majority of undergraduate students. This observation may have some implications for the effectiveness of WebCT and other Internet tools as a medium of communication.

The views being expressed in contributions reflected strongly held beliefs, or values, of the contributors. It was therefore very surprising to observe that contributors were so readily prepared to wear their hearts on the sleeves in the bulletin board. Especially given how reserved most of the participants tended to be in the face-to-face seminars. Presumably, the physical space that existed between participants also created some emotional distance between them and their posting to the bulletin board, which allowed the participants to open up to the class and so include more socio-emotional information in their postings. The group seemed to be acutely aware of how much some people were opening up to the group and couched their contributions to the board in ways that were supportive of other contributors.

The highly personal nature of the communication during the bulletin board meant that participants were required to communicate significant amounts of socio-emotional information (Lai 1998). Participants used the following five techniques during the bulletin board phase in order to enhance the sense of trust that had been developed during the course of the face-to-face seminar series.

- **Salutation;** participants tended to commence their contribution with a friendly, either personal or to the group, greeting of some kind.
- **Thanking;** participants thanked other members of the class for their pervious comments before moving onto making their own contributions to the discussion.
- **Personal references;** participants commonly referred to each other by use of their first names.
- **Acknowledging and praise;** participants often acknowledged and praised the contributions of other before making new contributions. Indeed, a small number of cases the contributions were solely comprised of acknowledgment and praise of other contributions.
- **Sharing personal information;** a number of participants really opened up to the class and prefaced their contributions by sharing a personal insight with the group.

At this stage it is interesting to note that at no time did any of the participants use emoticons to facilitate the discussion.

In order to enhance the trust that had already been developed within the class, participants tended to be courteous and uncritical, in a personal sense, but not an academic sense, of the contributions of others. That is, participants were careful to direct any critical or negative comments directly to the material posted on the bulletin board and not the author. This approach allowed participants to develop a degree of emotional distance between themselves and their postings. Such an approach also represents good practice in assignment marking. Experienced markers attempt to create some emotional distance between their students and their work so that students do not task negative feedback personally because feedback is directed at the work and not the student.. Contrary to what is often reported in the literature

concerning the conduct of participants in public discussion fora (Lai 1998). Participants did not use inflammatory comments. They also tended to use terms such as “personally” to indicate that their contribution was personal and a tentative view.

The participants also explicitly acknowledged the perspectives of the other participants by making comments such has “ I agree with what you are saying, but”. Indeed, it was clear that there was significant disagreement between participants. Nevertheless, the word “disagree” did not appear in the discussion. Again this observation does not accord with the literature, which tends to argue that heated debates prevail in computer-mediated discussions due to the impersonal nature of this medium (Lai 1998). However, it needs to be acknowledged that most studies that have been conducted into communication across networks have been conducted with undergraduate students. Perhaps in a postgraduate program where people have developed strong relationships as a result of spending three years studying together, which included eight months of face-to-face seminar, combined with the knowledge that the participants may work with each other at some time in the future, mitigated against the likelihood of a heated argument and personal attacks. It may also reflect the maturity and professionalism of the participants in this study. However, the most likely explanation of this paradox is that all the participants belonged to the same discourse community as a result of the length of time that they had spent studying at university and working in the education sector. The participants appeared to have been acculturated into the values and mores of academic argument, which facilitated vigorous debate about what the participants had to say, without degenerating into personal attacks.

The Chat Session

In stark contrast to the Bulletin Board, the Chat Session was a far more informal affair with the contributions being less well structured and much more immediate. Also in contrast to the bulletin board, the discussion did not address the issue we were asked to discuss. Indeed, there were two or three separate conversations going on that were totally unrelated to the set task. This situation was caused by the inability of the Chat Session facilitator to log onto WebCT and so facilitate the discussion. The discussion also tended to be dominated by people with good keyboard skills. For example, the author attempted to respond to a number of contributions. But by the time he had constructed his comments by hunting and pecking over the keyboard, the touch typists had moved onto to other points. The chat session was also interrupted for a protracted period as the WebCT server went down for quite a while.

Evaluation

Despite the apparent failure of the chat session four invaluable lessons emerged from the study. First, people who intend to use this technology need to ensure that everybody is adequately trained in the software to be used. Some people may benefit from additional training in keyboard skills. Second, it is important that participants confirm ahead of time that their hardware has sufficient power to operate these resource hungry packages. Third, the chat session needs to be effectively facilitated in order to ensure that contributors stick to the task at hand. Fourth, subject coordinators need to be aware of the strengths and weaknesses of bulletin boards and chat sessions. They both have different effects on the discussion in terms of the quality of the contributions to the discussion and the number of people who are included in the discussion. These differences can have a strong impact of the effectiveness of teaching activities and so subject coordinators need to be mindful of these differences when they decide to use the Internet as a teaching medium. Finally, it is vital that IT support is available immediately. Hence, facilitators need to work closely with IT support people to ensure that people are on hand if server or other network problems arise.

A CHECKLIST FOR ACTION

Bill Gates is often quoted as saying that "email is great to prepare for a meeting or to record a meeting, but it is not a substitute for a meeting". What Bill Gates is arguing here is that communication is a two way process and so is much more than the transfer of information between people. Nevertheless, information technology experts and educators are constantly looking for ways to use email and other Internet data transfer applications to create on-line classrooms. Real communication may not actually take place across the Internet, especially if users do not have good written communication or keyboard skills. Nevertheless, increased access to the Internet is set to play an important future role in reducing the transactional distance that adversely effects the educational performance of distance education students.

The basic reason for teaching across the Internet or any other computer network is to facilitate collaborative learning and not to deliver a subject in a predetermined and rigid format. In order for this goal to be achieved, Harasim *et al* (1995) argue that teachers should be mindful of the following points when using a bulletin board as part of delivering subjects over this medium:

- **Do not lecture:** An elaborate, logically coherent, but long sequence of comments often produces silence. Use short comments that are open-ended and which invite responses.
- **Clearly communicate your expectations:** Students need to be clear of the expectations that their teachers have of them for the subject as a whole and every component of the subject.
- **Be flexible and patient:** Teachers should guide the conversation and not dominate it.
- **Be responsive:** Ensure that every comment is responded to, especially at the beginning of an on-line subject or activity. If no one else replies, respond directly with a message or mention the author's comments in other messages.
- **Do not overload:** Teachers should contribute no more than one long comment each day, or even less if the students are actively contributing. Students are more likely to read and value several short notes, rather than one long entry.
- **Monitor and prompt for participation:** Teachers should frequently read the status reports provided by the system. Private messages should be sent to those students who are falling behind or are reading, but not writing. If students do not sign on for more than a week or do not reply to the private messages, then they should be telephoned and assisted with their problems.
- **For assignments, set up small groups and assign tasks to each group:** If the class is too large to have a single discussion space without overloading the students, then the class should be divided into two or more discussion groups.
- **Be a process facilitator:** Teachers should make sure that the students understand and abide by good 'netiquette' by not insulting each other or getting too far off the subject topic. Teachers should encourage meta-communication about the process, and make suggestions for improving the experience for all participants.
- **Write weaving comments:** Every week or so, teachers should write, or assign participants to write, comments that summarise the previous week's comments and focus the discussion.
- **Organise the interaction:** Electronic housekeeping includes moving or deleting items that do not belong in a particular conference and organising and modeling the use of key words and explicit references and associations among items to show relationships.
- **Set rules and standards for good netiquette;** and encourage meta-communication about anything that is causing the experience to be less valuable or enjoyable for all than it might be.

- **Establish clear norms for participation:** This includes developing procedures for grading on-line work that gives credit for good participation.
- **Allow others to play the role of facilitator:** Assign individuals or small groups to play the role of teacher and moderator for components of the subject.
- **Close and purge moribund conferences in stages:** Participants should be given an opportunity to save any messages that they wish to keep.
- **Adopt a flexible approach towards curriculum integration on global networks:** Curriculum areas will be affected by the diversity of opinions from different locations in the world. The best approach is to be open to changes and accepts new views on various topics.

Three of the four lessons learned during the course of this study also need to be added to this list.

- **Training:** People who intend to use this technology should be adequately trained in the software to be used. Some people may benefit from additional training in keyboard skills.
- **Hardware suitability:** It is important that participants confirm ahead of time that their hardware has sufficient capacity to operate these resource hungry packages and is compatible with the software to be used.
- **IT support:** IT support should be available immediately. Facilitators need to work closely with their support staff to ensure that support is available if the server or other network problems arise.

Internet teaching applications, such as WebCT, are increasingly being seen by educators, information technologists and administrators as a way of reducing the cost of external education that also improves the performance of distance students and expands the potential market size. However, the adoption of these applications requires significant investment by both students and universities in hardware and skills. These caveats need to be fully explored before any decision to teach across the Internet is made.

CONCLUSION

In a sense, undergraduate students can be thought of as belong to a variety of different cultural groups of tribes (Beasley 1987; 29) because they have yet to be acculturated into the value systems and norms of the university. The cultural differences that exist between the university and its students can create barriers to communication, leading to the angry exchanges that have been identified in the literature. However, as noted above, even though the participants in the cyber colloquium were drawn from a wide variety of backgrounds, they were all drawn for the same discourse community (Kokkinn, Head, Feast, and Barrett 1998). Hence, they understand the norms and values of the university. Moreover, they understand the rules of engagement during an academic argument. In particular, it is vital that socio-emotional space be created between students and their postings. Or to use a football metaphor, you play the ball not the man. Consequently, criticisms of a piece of work are unlikely to be taken personally, so angry exchanges are unlikely to take place.

These observations point to the need for commencing university students to adapt to the distinctive ways of university behaviour and its values (Drury and Webb 1991). This adaptation, or more correctly acculturation, can be described as a period of apprenticeship, in which students adapt to the specific disciplinary discourse and knowledge (Cope and Kalantzis 1993). However, the apprenticeship is not a formal or structured process. Rather it is an unstructured process in which university students respond to the explicit, but much more commonly the implicitly, cues from their lecturers and fellow students. Unfortunately, distance education students often have little if any contact with their lecturers and fellow students. Hence, this process of acculturation is unlikely to occur. It is however, vital that

this process does occur if communication across the Internet is to be effective. There are probably two approaches that can be used to facilitate the acculturation process. The first is for distance education courses to include a residential component that focuses on this process of acculturation. The second is for activities that facilitate this process to be embedded into distance education materials. Whatever approach is used, the key point for students to appreciate is that a certain amount of socio-emotional distance needs to be created between students and their work and that the focus of any debate about the work of students should be directed solely at the work and not be personalised.

REFERENCES

Beasley, V. (1987) *Eureka! Or How to be a Successful Students*, Flinders University, Adelaide.

Caladine, R. (1993) *Overseas experience in non-traditional modes of delivery in higher education using state-of-the-art technologies: A literature review*. Canberra: Department of Employment, Education and Training, Occasional Paper Series.

Cope, B. and Kalantzis, M. (eds.) (1993). *The Powers of Literacy*. London: The Falmer Press.

Department of Education, Training and Employment (1998) *Community Partnerships in Education 1998, Volume 1*. Adelaide: Department of Education, Training and Employment.

Drury, H. & Webb, C. (1991). Literacy at Tertiary Level: Making explicit the writing requirements of a new culture. In F. Christie (Ed) *Literacy in Social Processes: Papers from the Inaugural Systemic Linguistics Conference*. Darwin: Centre for Studies of Language Education, Northern Territory University, 214-227.

Gengler, B. (2002) That :-) whimsical aside turns 20, *The Australian IT*, 24 September 2002, 3.

Harasim, L. (ed.) (1995) *Global Networks: computers and International Communication*. Cambridge. Massachusetts: The MIT Press.

Harasim, L., Hiltz, S.R., Teles, L. and Turoff, M. (1995) *Learning Networks: A Field Guide to Teaching and Learning Online*. Cambridge. Massachusetts: The MIT Press.

Jones, D. (1996) Computing by distance education: Problems and solutions. Proceedings of the ACM SIGCSE/SIGCUE Conference on Integrating Technology into Computer Science Education, Barcelona

Kokkinn, B. Head, M., Feast, V. and Barrett, S.R.F (1998) Transforming the Teaching of Economics: Embedding Tertiary Literacy. Paper presented at the HERDSA Conference98, Auckland University , July 1998.

Lai, K-W. (1998) Social interaction and communication a web-based tertiary course: Some observations, *Proceedings of the Sixth International Conference on Computers in Education*. Beijing, China, Volume 1, 79-85.

Computer Adventure Games as Problem-Solving Environments

David D Curtis
School of Education Flinders University *david.curtis@flinders.edu.au*

Michael J Lawson
School of Education Flinders University *mike.lawson@flinders.edu.au*

Claims that computer-based adventure games are productive environments for the development of general problem-solving ability were tested in a study of 40 students' interactions with a novel computer-based adventure game. Two sets of factors that are thought to influence problem-solving performance were identified in the literature – domain-specific knowledge (schema) and general problem-solving strategies. Measures of both domain-specific knowledge and general strategy use were developed and applied in the study. A cognitive model to explain performance is developed in which there are complex relationships among key concepts. General strategies were found to have important influences on problem-solving performance, but schema was negatively related to performance. The implications of these findings for both classroom practice and future research designs are discussed.

Computer adventure games, problem-solving ability, strategies

INTRODUCTION

The claims made about benefits that students might derive from the use of adventure games can be categorised into three groups: those about social interaction, about language, and about problem-solving. Several authors have suggested that children develop greater social skills through their interactions while playing adventure games (Craig, Podmore, Atmore, & Ashworth, 1987; Heron, 1987; Sherwood, 1988). Sherwood (1988, p.299) noted that: "The computer fosters cooperative work among students as the goal is not necessarily to complete but to problem solve. Students who collaborated on the adventure game often taught each other. Such computer oriented group work provided a positive socialising experience."

Others have claimed, on the basis of observational studies, that students develop enhanced language skills (Rice, 1985; Sherwood, 1988; Strack, 1985; Thompson & Duncan, 1988; Unwin, 1983). For example Unwin (1983, p.149) argued that "As a young person plays an adventure game, he develops such important skills as spelling, reading comprehension, critical thinking, and creativity".

It has been argued that adventure games often include puzzles that need to be solved in order to complete the game (Heron, 1987); that students need to apply knowledge derived from "life and literature" (Rice, 1985); that students must use general problem-solving skills like inferring, monitoring, and deductive reasoning (Sherwood, 1988); and that adventure games encourage application of metacognitive skill (Henderson-Lancett & Boesen, 1986). For these reasons, adventure games are claimed to have application across a range of curriculum areas (Bell & Scott, 1988; McArdle, 1985). Taken together, this final group of assertions

support the view that students become better problem-solvers following exposure to adventure games.

The focus of this study was the claims for enhanced problem-solving performance flowing from the use of computer-based adventure games. Many of the studies cited above are anecdotal reports of observations made by teachers and others in classrooms as children used adventure games. In one of the few empirical studies located in a review of the literature, Grundy (1988) found that adventure games do have potential as effective problem-solving environments, but that this potential is not realised because children are often able to invoke techniques for avoiding the use of transferable strategies. Grundy noted that in many adventure games children are able to avoid reading for detail and are often not required to assimilate new information for later recall and use. The results of Rhodes' (1986) study also cast doubt on some of the claims made for adventure game use. Rhodes found no differences in comprehension skills between a group of students who used adventure games and a control group. Classroom use of many adventure games for influencing problem-solving is also limited because the programs usually do not include guidelines to suggest how they might be used, what problem-solving strategies could be developed, or what other experiences might be arranged to support their use (Grundy, 1988, p.21). There is, therefore, uncertainty about the status of the view that use of adventure games will result in improved problem-solving performance. The research basis for this claim is not extensive and is generally not embedded in a suitably developed conceptual framework that would provide the background for relating particular features of adventure game use and processes invoked during problem-solving.

COGNITIVE PERSPECTIVES ON PROBLEM-SOLVING

Problems and Problem-Solving

Mayer (1992) presented a useful definition of problem-solving:

> *Problem solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver.*

Greeno, Collins and Resnick (1996) outlined three major theoretical stances on learning and problem-solving that they termed associationist/behaviourist/empiricist, cognitive/rationalist, and pragmatist/situative/sociohistoric. Of these, the variants of the cognitive and situative appear to be most useful in developing understandings of individuals' development of problem-solving capability.

In the information processing framework for problem-solving developed by Newell and Simon (1972), a problem exists when the current arrangement of problem elements (the current state) is different from the desired arrangement (the goal state) and operators that can effect the transition between the two states are not readily available. The problem solver's task is to find the operators that will enable the current state to be transformed into the goal state. The states of an adventure game are the set of locations and their descriptions, the set of objects and their positions and functions, and the set of restrictions on what actions can be taken. The problem operators are the commands that players issue to effect a change in the game state to bring it closer to the goal. Using the Newell and Simon model, the strategic nature of the student's actions (moves) is likely to have a major influence on the success of the problem-solving attempt.

However, more recent research has also highlighted the important influence that the student's store of knowledge has on comprehending the situation, selecting moves, and thereby affecting the outcome (Schneider, 1987, 1990).

The strictly cognitive models of problem-solving, with an emphasis on abstract representations inside the problem-solver's head, is criticised by researchers who have investigated problem-solving in 'real world' situations. Scribner (1986) described the processes used by packers in a warehouse and Lave (1988) showed how shoppers made decisions about which products, packaged in different amounts, represented 'best buys'. The individuals in these studies performed skilfully, but did not use abstract processes. Rather, they manipulated concrete materials, used non-formal practical representations, and employed low mental load strategies.

Problem-solving ability is defined as "cognitive processing directed at achieving a goal when no solution method is obvious to the problem-solver" (Mayer & Wittrock, 1996). This definition emphasises neither general processes nor the individual's knowledge base. It may be taken to include either or both in effective problem-solving performance.

Expertise, Knowledge, and Strategy Use in Problem-Solving Performance

Chi, Glaser, and Rees (1982) claimed that success in problem-solving depends on access to a well-developed and extensive knowledge base. Experts have richly elaborated knowledge bases and they use these early in problem-solving to develop a representation of the problem task that then directs the moves used to solve the problem. Thus, Chi et al. (1982) argued that domain specific knowledge is of major importance in expert problem-solving. Others also support this view (Larkin, 1985; Schneider, 1987). Sweller (1990) went further and argued that:

> Subsequent work on expertise in areas such as physics and mathematics supported the suggestion that domain-specific-knowledge rather than general problem-solving skills differentiated novices from experts. (Sweller, 1990, p.412)

Sweller postulated that experience in a domain results in the formation of a schema for that domain that includes a knowledge base and a set of highly automated rules. The schema enables problem classification according to previously encountered solution procedures and the associated rules can then be used to direct performance. If this is so, schematic knowledge should be strongly associated with a high level of success in problem-solving.

Other researchers have argued that it is necessary to qualify this view of the dominance of prior knowledge as a factor in problem-solving performance. They have argued that, through experience in solving problems, students also use and develop a set of more general skills that they may apply in new situations. Bereiter and Scardamalia (1986) pondered the question of how novices, with limited domain-specific knowledge, transform themselves into experts. They argued that some people are expert at becoming expert and that they do this by the use of strategies. This view is taken further, with claims that students can be taught to use a set of general problem-solving skills (in a domain), and that when the skills are well developed, students will transfer them to other domains. A number of authors have reported improved performance following strategy instruction (Bereiter & Scardamalia, 1987; Charles & Lester, 1984; Clements, 1990; Hembree, 1992; Lawson & Rice, 1987; Paris, Wixson, & Palincsar, 1986). There is, however, more argument about the extent to which students can spontaneously transfer strategies developed in one area to another, so that even when subjects have knowledge available, frequently they only apply that knowledge when reminded of its availability and relevance (Gick & Holyoak, 1983; Ross, Ryan, & Tenpenny, 1989).

Thus there are two different positions: one that holds that effective problem-solving depends upon a well developed knowledge base and another that relies upon generally applicable strategies. Despite the identification of these two contrasting positions it is not necessary, or

even helpful, to opt for only one or other of schema induction or strategy use to explain problem-solving performance. Siegler (1990) argued strongly that schematic knowledge and strategy use interact and that this interaction requires acknowledgment of the contribution of both sets of factors to the students' outcomes. In the initial acquisition of knowledge, strategy use is important, and in later access to and use of that knowledge, it is again a factor (Alexander & Judy, 1988; Chi, Hutchinson, & Robin, 1989; Chi & VanLehn, 1991; Prawat, 1989).

Thus the literature suggests that there are two major sets of influences on problem-solving, schematic knowledge and strategy use, that they interact, and that they need to be considered in explanations of problem-solving performance. However, schematic and strategic knowledge cannot be the elemental determinants of performance. Schematic knowledge must arise from experience and also be related to ability. Strategy use is very likely to be related to ability and may also depend upon experience.

In the adventure game context it was hypothesised that both general ability and experience of adventure games would enhance problem-solving skill, and from this it was predicted that more adventure game experience would result in superior adventure game performance. It was also postulated that experience in use of adventure games would affect students' knowledge of these games and their use of general problem-solving strategies. These two factors, separately or in concert, could lead to enhanced performance. These assumptions are summarised as a set of hypotheses:

1. Higher levels of ability will lead to enhanced performance;
2. More experience of adventure games will lead to enhanced performance;
3. Higher ability will be associated with greater general strategy use;
4. More experience of adventure games will lead to greater schematic knowledge;
5. Greater use of general problem-solving strategies will lead to superior performance;
6. A greater domain-specific (schematic) knowledge base will lead to superior performance.

These relationships are summarised in Figure 1. The two independent variables (Ability and Experience), the two mediating variables (Strategy use and Schema), and the dependent variable (Performance) are all latent constructs. They are not measured directly but are indicated by a range of manifest variables. For clarity, the manifest variables that are used to operationalise the constructs have not been included in the diagram.

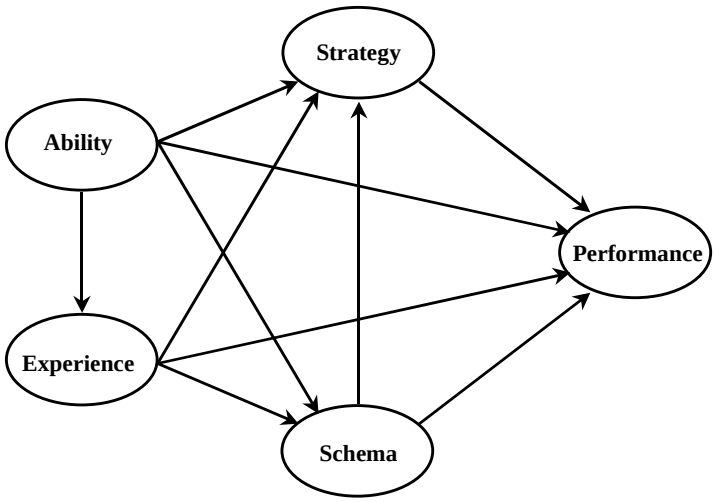


Figure 1. The hypothesised model of relationships among Ability, Experience of computers and adventure games, Schema for adventure games, Strategy use and Performance on adventure games

In the model, some relationships in addition to those specifically posited in the hypotheses were tested. A path was included from Ability to Experience in order to detect whether there was a bias in the inclination of students of different ability levels to engage in the use of computers or adventure games. Two other paths have been postulated in the model to be tested. One is a path from Ability to Schema. The presence of this path would indicate that higher ability students induce adventure game schema more readily than students of lesser ability on the basis of a given amount of adventure game experience. Similarly, a path between Experience and Strategy use is proposed. Here the main hypothesis is that Ability is related to Strategy use, but that more experience of adventure games may result in greater application of general problem-solving strategies. This possibility is also tested by the inclusion of a path from Schema to Strategy use. The import of this path is that the influence of Experience may operate directly on Strategy use, but it may also operate indirectly through Schema.

The path model includes the possibilities that there are direct relationships between constructs, for example from Ability to Performance, but that there may also be indirect relationships, for example Ability may also influence Performance through Strategy use.

Thus, a model of problem-solving performance in adventure games was hypothesised that invoked strategy use and schematic knowledge as mediating constructs with cognitive ability and domain-specific experience as primary constructs. The ways in which these constructs were operationalised are described in the Method section of this paper, and the corresponding variable labels are shown in italics.

METHOD

Subjects

Participants in the study were 44 students from three metropolitan schools in Adelaide. The students ranged in age from 12 to 15 years and they were in Years 7 to 10. Data from four students were lost due to equipment failure, leaving complete data for 40 subjects, of whom 18 were female.

Prior Knowledge and General Ability

In order to assess the extent of previous exposure to adventure games, students completed a self-report questionnaire in which they were asked about their experience of, and affect for, computers (*CmpExp* and *Affect*), and the number of adventure games that they had played (*AgExp*). The questionnaire also sought information on students' age (*Age*), sex (*Sex*) and knowledge of meanings for 20 words that were taken from the adventure game (*AWK*). In order to generate information about students' prior schematic knowledge of adventure games, the questionnaire also presented a location description from another adventure game as it might have appeared on screen. Students were asked to generate a list of moves that they thought would be appropriate (*Elab*), and then to select the move that they thought would be the best one (*Action*). This list of moves and the suggested best move were scored to establish a measure of schematic knowledge for adventure games.[1]

Students were classified as novice, intermediate or experienced players based upon the number of adventure games that they indicated they had played. Those who had played none or one game were classified as novices, of whom there were 10; those who had played two

[1] All material used in this study, including the adventure game program, are available from the first author

or three games as intermediate, of whom there were 20; and those who had played four or more games as experienced, and there were 10 in this category.

Students also completed a standard word knowledge test (Australian Council for Educational Research, 1989) to provide an estimate of verbal ability (*WKTF*). This was taken as a proxy for general ability.

At the conclusion of the session in which students completed the questionnaire and the word knowledge tests, each student was given a copy of the instructions for playing the adventure game. They were asked to read the instructions before presenting for the adventure game session which was held one week later.

The Adventure Game

For this study a new text based adventure game, The Ancient Abbey, was developed. It has 34 locations and six objects that must be found in order to complete the game. It is structured so that the first two objects can be found simply by going to their initial locations. Both must be found to retrieve the third object, and this is required to move to the section of the game where the remaining objects are placed. In this area, locating the sixth object requires the player to hold the fifth, and to get this, object four must already have been obtained. In this way, the number of objects located and the number of game locations visited provide an index of performance in the game.

The program includes data collection code. As the player makes a move, the time of the move, the player's location, and the command issued are recorded in a text file.

Think-Aloud Protocol Generation Training

Immediately prior to the adventure game session, students were given training individually in the generation of a concurrent think-aloud protocol. In the training, which used another text-based adventure game with the same screen layout as the experimental game, the researcher modelled the think-aloud process by reading relevant information from the screen, by articulating possible moves, and by selecting and giving reasons for the chosen move. Students were then asked to continue with the game. During the training, if students did not give reasons for their moves, they were reminded to do so, and in some cases the process was modelled again. The duration of this training was approximately 15 minutes, but varied depending upon the extent to which students demonstrated their ability to articulate reasons for their moves. The experimental game was commenced when the researcher was satisfied that the participant would be able to generate a useful protocol.

Adventure Game Data Collection

When students presented for the adventure game session, they were given a copy of the adventure game instructions and a blank 'map' (a sheet of paper with a grid of boxes) that could be used to keep records of their progress. Students were not told to use this, it was simply available for those who chose to use it. Further, while students played the game, they were reminded of the need to provide a concurrent think-aloud protocol. The think-aloud protocols were audio-tape recorded and later transcribed. The transcripts were analysed for evidence of students' use of a variety of strategies, including game specific strategies such as saving the game if they thought that the next move might be dangerous, and more general strategies like making inferences from information presented on screen, planning a sequence of moves, or monitoring performance. The strategies coded in the analysis of transcripts are shown in Table 1.

A criterion measure of student performance in the game was required. The adventure game is in effect a sequence of small puzzles. In choosing to move from one location to the next, there is a description of the location and an indication of possible directions. Using the hints provided in the location descriptions helps users to decide courses of action. In order to complete the game, six objects had to be collected. The first two could be found by visiting their locations and issuing the command to take them. However, one of the objects was a key that had to be used to open a door to a location where other objects were available. Thus the number of objects collected reflected the degree of completeness of the game solution. Thus the two variables, *Locatns* and *Objects*, were used as indicators of performance. Other information collected during students' use of the game were the extent to which they used maps (*Maps*) and the number of moves that they made (*Moves*).

Table 1. Summary of strategic behaviours coded in transcripts of students' protocols recorded during adventure game play

	Planning
P1	Single moves with no forward planning indicated.
P2	A move for which a goal is specified.
P3	An adventure specific move like saving game status.
	Recognition
R	Recognition of the significance of an object.
	Inference
I1	Infers use for an object when it is first encountered.
I2	Infers use for an object when prompted by a situation.
	Monitoring
M1	Recall that a location has been seen before.
M2	Acknowledgment of a memory limitation.
M3	Acknowledgment that an error has been made.
M4	Summary reflection on performance.
	Errors
E1	Error in reading text from screen.
E2	Error in recall of a location.
E3	Error in using a direction. Says East but moves West.

Data Analysis

Two approaches were taken in the analysis of the data collected during this investigation. First, in order to ensure that the data were free of errors, a range of descriptive statistics were computed. The hypothesised relationships among the variables were investigated using *t* tests to compare independent samples. Although this method provides a useful exploratory approach to data analysis, it lacks the power to investigate a problem that is multivariate in nature. Second, path modelling was undertaken in order to establish a more complete understanding of the explanatory constructs invoked in the study, and the focus in the remainder of the paper is on the conduct and interpretation of the path analysis.

In order to further examine the relationships among the set of variables, the path model shown in Figure 1 was tested using Partial Least Squares (PLS) path analysis with the program PLS Path (Sellin, 1987). Alternatives available for this analysis included multiple regression analyses and the use of Structural Equation Modelling (SEM). The former was tried but rejected because of the detection of multicollinearity among independent variables. The latter method requires large sample sizes and assumes multivariate normal distributions among variables. Given that the sample size was relatively small at 40, PLS was the chosen analytical method. Because it does not make distributional assumptions, PLS has the disadvantage that standard errors are not computed from a distribution. However, estimates of standard errors are provided through the use of the jack-knifing technique, and so it is possible to infer the significance of the variables included in the model. Path analysis using

PLS shares with SEM the capacity to define the problem in terms of latent constructs and manifest variables and to explore theoretically interesting relationships among the latent constructs.

RESULTS

Initially a path model was specified that included all possible paths among latent variables as shown in Figure 1. From two to eleven manifest variables were used as indicators for each latent variable. In an iterative refinement of the model, some manifest variables were removed from the model when their loadings were found to be low compared with the jack-knife standard error of the estimate. Some paths between latent variables were also removed when their magnitudes were low compared with the estimated standard errors. The criterion for retention for both manifest variables and for paths was that the estimated magnitude of the manifest variable loading or the path coefficient had to be more than twice the jack-knife standard error.

During the refinement of the model five manifest variables were removed from the Strategy use latent construct. The variables removed included the three behaviours coded as Errors. These involved errors in reading information (*E1*), recall of information (*E2*), and errors in making moves (*E3*). Two monitoring behaviours were also deleted from the model. These were *M1*, recall of a location, and *M2*, recognition of memory limitation. Another variable of interest, *P1* – making moves without articulating a purpose – had not been included in the model as this variable represents a lack of strategic activity rather than the use of a particular strategy.

Several paths between latent constructs were also removed because they proved to be non-significant. A path had been hypothesised between Ability and Experience. This had been included in case there had been bias in that high ability students might have been more (or less) inclined to play these games or otherwise to engage with computers. This path coefficient was 0.156 with a jack-knife standard error of 0.178, and so it was removed. It can therefore be concluded that there was no significant ability bias in the tendency to engage with computers or to play adventure games. The lack of a relationship between these variables enables them to be treated as independent exogenous constructs in the model.

The path from Experience to Strategy use was removed initially. The path coefficient was 0.125 and the jack-knife standard error was 0.180, indicating that the path must be regarded as non-significant. However, following the removal of the path from Schema to Strategy use (see below), this path was reinstated and its magnitude was marginally significant at 0.218 with a standard error of 0.108, the criterion ratio being 2.180.

A path had been postulated from Schema to Strategy use. It was hypothesised that having a schema for adventure games might enable the selection and deployment of particular strategies. However, the path coefficient was 0.190 with a jack-knife standard error of 0.230. While this path is of moderate magnitude, the relatively high standard error may indicate that more experienced players with better developed schemas use them directly in their solutions, while less experienced players with limited schematic knowledge may rely largely upon general purpose strategies. The relatively high standard error estimate of this path coefficient may be important in understanding the relationships that emerged in the analysis, and this matter is again raised in the Discussion section.

The results of this analysis following refinement are summarised in Figure 2.

Table 2 shows the results of the outer model – the loadings of manifest variables on their associated latent variables, with the jack-knife standard errors of those loadings and the

ratios of the loadings to the standard errors. These ratios provide an indication of the consistency the loadings. For most latent variables, the sets of loadings are reasonably high, but for the Strategy use latent variable, the loadings are somewhat variable and the Q^2 statistic is rather low. This may suggest that there is considerable variation among participants in the particular strategies that they use.

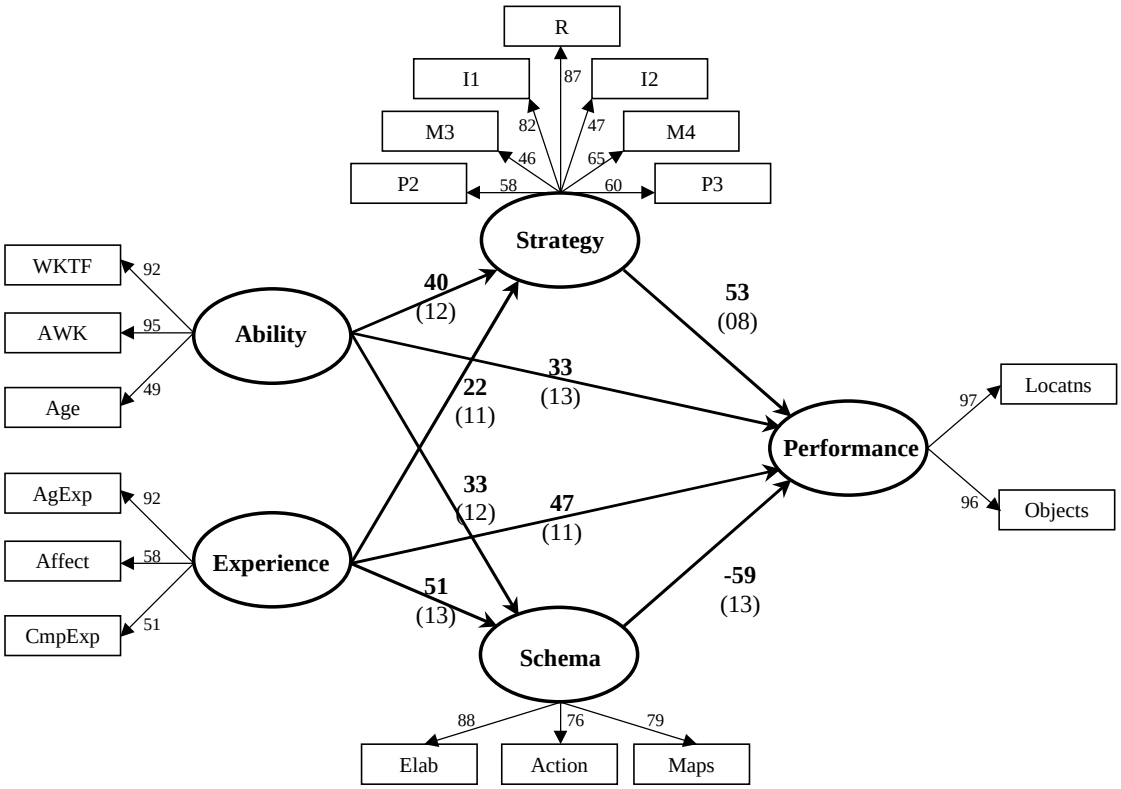


Figure 2. Results of a path analysis on the hypothesised model of adventure game performance (Coefficients are shown for inner model paths with standard errors in parentheses. Outer model loadings are shown without standard errors, given in Table 2).

Table 2. Loadings of Manifest Variables on Latent Variables in the Path Model

Latent Variable	Manifest Variable	Loading	Jack-knife Standard Error	Ratio of Loading to Standard Error
Ability	WKTF	0.921	0.034	27.088
	AWK	0.946	0.023	41.130
	Age	0.486	0.122	3.984
Experience	AgExp	0.919	0.024	38.292
	Affect	0.581	0.109	5.330
	CmpExp	0.513	0.115	4.461
Schema	Elab	0.879	0.040	21.975
$Q^2 = 0.303$	Action	0.763	0.069	11.058
	Maps	0.791	0.065	12.169
Strategy use	P2	0.542	0.189	2.868
$Q^2 = 0.077$	P3	0.532	0.148	3.595
	I1	0.787	0.111	7.090
	I2	0.473	0.126	3.754
	M3	0.526	0.129	4.078
	M4	0.695	0.270	2.574
	R	0.887	0.077	11.519
Performance	Locatns	0.966	0.009	107.333
$Q^2 = 0.459$	Objects	0.956	0.011	86.909

(The Q^2 statistic is an indication of the predictive power of predictor variables for the manifest variables associated with a given latent variable).

The results of the inner model, the paths among latent variables, are shown in Table 3. In each cell the path coefficients between latent variables, their jack-knife standard errors, and their ratios are tabulated. Also shown are the R^2 values for the predicted variables. The R^2 value for Performance is 0.590 indicating that the model is able to predict a high proportion of the variance in this construct.

Table 3. Estimated Path Coefficients, Jack-knife Standard Errors and Criterion Ratios in the Refined Model of Performance

		Endogenous (Predicted) Latent Variables		
		Schema $R^2 = 0.403$	**Strategy use** $R^2 = 0.227$	**Performance** $R^2 = 0.590$
Predictor Latent Variables	Ability	0.328 0.120 2.733	0.403 0.117 3.444	0.330 0.126 2.619
	Experience	0.513 0.131 3.916	0.218 0.108 2.019	0.473 0.111 4.261
	Schema			-0.590 0.128 -4.609
	Strategy use			0.534 0.080 6.675

Each cell has the path coefficient (Beta), the Jack-knife standard error, and their ratio

Table 4 shows the direct, indirect and total effects of the explanatory latent constructs on dependent ones. Both Ability and Experience are mediated through Strategy use and Schema. Ability has positive direct and indirect effects, while Experience has a positive direct effect, a small positive indirect effect through Strategy use, and a substantial negative indirect effect through Schema.

Table 4. Summary of Path Coefficients among Latent Variables

	Direct Effect	**Indirect Effect**	**Total Effect**
Schema			
Ability	0.329	.	0.329
Experience	0.513	.	0.513
Strategy use			
Ability	0.403	.	0.403
Experience	0.218	.	0.218
Performance			
Ability	0.329	0.212	0.351
Experience	0.473	-0.187	0.287
Schema	-0.590	.	-0.590
Strategy use	0.534	.	0.534

Discussion of Results

The positive relationships from Ability to Performance and Experience to Performance are expected. Similarly the positive paths from Ability to Strategy use, from Ability to Schema, from Experience to Schema and from Experience to Strategy use are also expected. The

relative magnitudes of these paths are also of interest. Ability has a moderately strong path to Strategy use and Experience has a similarly strong path to Schema.

Of greater interest are the influences of the modelled constructs on Performance. Predictably, both Ability and Experience have positive direct influences on Performance. The quite strong positive path from Strategy use to Performance indicates that the use of general problem-solving strategies has been an important element in the mechanisms by which the participants in this study have been able to solve the problems posed within the adventure game environment. However, what is quite surprising is the strong negative path from Schema to Performance. At first glance, this result suggests that having a well developed Schema for adventure games is counter-productive to achievement, a proposition that makes no theoretical or practical sense. Even if an extreme information processing position was taken and it was asserted that problem-solving performance was a result only of the application of very general strategies, how could the acquisition of greater knowledge undermine performance?

In order to understand the unexpected relationship between Schema and Performance, the measures for Schema must first be reviewed. The manifest variables used to form the Schema latent construct in the study may not be as complete a representation of this construct as is desirable. Schema for adventure games was inferred from a scenario taken from another adventure game from which participants had been asked to generate a range of moves and from them to select optimal moves. These two variables are taken as indicators of an adventure games schema: those participants who had a well developed schema for adventure games were presumed to be capable of generating a greater number of moves and of selecting more productive moves. This appears to provide an indication of schematic knowledge, but it may not provide as complete a representation of it as is desirable. More detailed questions about possible courses of actions and expected outcomes may well have provided a more comprehensive indicator of this latent construct.

Two explanations are hypothesised for the observed negative path between Schema and Performance. First, it was assumed that the students participating in this study represented a range of experience from novice to some advanced level of expertise and that there would be a gradual transition in the use of knowledge and strategies with increasing experience. However, it is possible, that as experience grows, there is a change in the architecture of problem-solving and that a model that describes near-novice problem-solving is rather different from the model required to account for proficient or expert problem-solving. If this is so, then in the current study the hypothesised model is a conflation of two, and possibly more, models. There are indications in the current model that this may be the case. Two paths involving Schema, one from Ability and one to Strategy use, had rather high standard errors. In addition, the Q^2 statistic for the Strategy use latent variable was rather low suggesting that strategy use among participants was variable. It is possible that the very inexperienced participants, having no prior experience of adventure games, were forced to rely only on whatever general strategies they could to solve the puzzles that were part of the game and to develop an understanding of this particular game context from the descriptions provided. Those with more experience may have been able to use their knowledge base to make better strategy choices. This line of argument helps to explain the positive path from Experience to Strategy use and from there to Performance. It also suggests that for more experienced individuals, there should be a significant path from Schema to Strategy use. However, it does not explain the negative path from Schema to Performance.

A second related hypothesis is required. With more experience, schematic knowledge develops, and that knowledge enables the selection of efficient strategies. However, despite the range of experience reported among participants in the present study, it is suggested that

their experience ranged from novice to perhaps competent, but that it did not include experts. The level of schematic knowledge available to the competent participants enabled them to make more effective use of general strategies, but was not automated to the extent that it could drive performance directly. Thus, in the composite model, those participants with better developed schemas used them to employ general strategies more effectively. Thus, for all levels of experience among participants there is a positive association between Strategy use and Performance. For complete novices, there is little schematic knowledge, while for their slightly more experienced peers, there is greater schematic knowledge but even greater strategy use. Thus, greater schematic knowledge is negatively associated with Performance.

Somewhat paradoxically, these hypotheses may support the views of Sweller (1999) that informational complexity and the application of general problem-solving strategies place a high cognitive load on individuals. The developing knowledge base may reduce the cognitive load associated with using the general strategies required to apprehend and understand the virtual environment being encountered and enable efficient use of those strategies during the problem-solving of non-experts.

These hypotheses have implications for research designed to elicit cognitive models of performance at various stages of the transition from novice to expert.

CONCLUSION

This study was designed to test claims that experience of adventure games leads to the development of general problem-solving skills. From the literature on problem-solving two possible factors, schema and strategy use, were identified that could influence problem-solving performance. Evidence for effects of adventure game experience on both schema and on strategy use were sought, and a model of their influence on performance was developed and tested.

The model of adventure game performance based on data collected in the study revealed that experience of adventure games leads to a modest increase in the use of general strategies and that the application of those strategies does lead to enhanced performance in a novel adventure game task.

However, it is not possible to assert that these strategies will be used by individuals when confronted with new problems in other domains. On this basis, it would be imprudent to advocate the use of adventure games in classrooms as vehicles for the development of broadly applicable problem-solving ability. Additional work is required to investigate whether the enhanced use of general strategies does transfer from adventure game environments to other problem domains that students might encounter in their classrooms and beyond.

In the past, general strategy use and schema were offered as alternative explanations for problem-solving performance, and a crude interpretation of the model developed in this study might be used to assert a case for the predominance of general strategy use in solving adventure game problems. However, it was argued in the discussion of the results that the model developed here may be only one of several that could be invoked to explain performance of individuals as they move from novice status through competence to expert. Thus no conclusion is suggested for the relative importance of general strategies and schematic knowledge. Instead, Siegler's advice is reiterated when he sought to show:

> ... how specific knowledge influences choices among strategies, how choices among strategies in turn influence the construction of specific knowledge, and how individual differences in both

> initial knowledge and cognitive style influence both choices among strategies and acquisition of further knowledge. (Siegler, 1990, p.74)

Thus, Seigler supports both the complementary roles of general strategy use and schematic knowledge and the notion that cognitive mechanisms may change over time reflecting growth in both schematic knowledge and the availability of general strategies.

It is suggested that further research is required in order to establish whether different cognitive mechanisms are associated with the various stages between novice and expert status. If individuals can be assigned to groups on this continuum, it may be possible to construct path models for each group and to compare the models. Alternatively, a longitudinal study may track individuals as their expertise develops over time within a domain.

REFERENCES

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*(4), 375-404.

Australian Council for Educational Research. (1989). *ACER Word Knowledge Test. Form F*. Hawthorn: ACER.

Bell, S., & Scott, I. (1988). *Springboards: ideas for using computers in the classroom*. Melbourne: Nelson.

Bereiter, C., & Scardamalia, M. (1986). Educational relevance of the study of expertise. *Interchange, 17*(2), 10-19.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, N.J.: Lawrence Erlbaum.

Charles, R. I., & Lester, F. K. (1984). An evaluation of a process oriented instructional program in mathematical problem solving in grades 5 and 7. *Journal of Research in Mathematics Education, 15*(1), 15-34.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, N.J.: Lawrence Erlbaum.

Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly, 35*(1), 27-62.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *The Journal of the Learning Sciences, 1*(1), 69-105.

Clements, D. H. (1990). Metacomponential development in a Logo programming environment. *Journal of Educational Psychology, 82*(1), 141-149.

Craig, B., Podmore, V., Atmore, D., & Ashworth, D. (1987, December). *Computers in the classroom: observations of children's social behaviour and collaborative group work.* Paper presented at the Australian Association for Research in Education / New Zealand Association for Research in Education Conference, Christchurch, NZ.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1-38.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.

Grundy, S. (1988). *The computer and the classroom: critical perspectives.* Paper presented at the Educational research: Indigenous or exotic? Annual Conference of the AARE., Armidale, NSW.

Hembree, R. (1992). Experiments and relational studies in problem solving: a meta-analysis. *Journal for Research in Mathematics Education, 23*(3), 242-273.

Henderson-Lancett, L., & Boesen, J. (1986). Dragon World 1: implementing an adventure game. In A. D. Salvas & C. Dowling (Eds.), *Computers in education: The crest of a wave* (pp. 115-117).

Heron, J. (1987). Dread Dragon Droom. *Classroom Computing, 7*(1), 15-23.

Larkin, J. H. (1985). Understanding, problem representations, and skill in Physics. In S. F. Chipman, J. W. Segal & R. Glaser (Eds.), *Research and open questions* (Vol. 2).

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, N.Y.: Cambridge University Press.

Lawson, M. J., & Rice, D. N. (1987). Thinking aloud: analysing students' mathematics performance. *School Psychology International, 8*(4), 233-244.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W H Freeman.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47-62). New York: Macmillan.

McArdle, E. (1985). Adventure games in the classroom. *Com 3, 11*(3), 13-14.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice Hall.

Paris, S. G., Wixson, K. K., & Palincsar, A. S. (1986). Instructional approaches to reading comprehension. In E. Rothkopf (Ed.), *Review of research in education*. Washington, D.C.: American Educational Research Association.

Prawat, R. S. (1989). Promoting access to knowledge, strategy, and disposition in students: a research synthesis. *Review of Education, 59*(1), 1-41.

Rhodes, J. (1986). Adventure games in the classroom: a case study. In B. Rasmussen (Ed.), *The information edge: The future for educational computing* (pp. 111-116).

Rice, S. (1985). Adventure games - computers and language. *Classroom Computing, 4*(3), 10.

Ross, B. H., Ryan, W. J., & Tenpenny, P. L. (1989). The access of relevant information for solving problems. *Memory and Cognition, 17*(5), 639-651.

Schneider, W. (1987). *The knowledge base and memory performance: a comparison of academically successful learners.*

Schneider, W. (1990, April 16-20). *Domain specific knowledge and cognitive performance.* Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Scribner, S. (1986). Thinking in action: some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence. Nature and origins of competence in the everyday world* (pp. 13-30). Cambridge: Cambridge University Press.

Sellin, N. (1987). PLS Path (Version 1.8) [Statistical analysis software]. Hamburg: Sellin, N.

Sherwood, C. (1988, September). *Adventure games. A golden opportunity for young learners. The impact of computers on the learning of young children.* Paper presented at the Golden opportunities: Sixth Australian Computers in Education Conference, Mount Lawley, Perth, WA.

Siegler, R. S. (1990). How content knowledge, strategies, and individual differences interact to produce strategy choices. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance*. New York: Springer-Verlag.

Strack, G. (1985). Expedition to Saqqara. *Classroom Computing, 5*(2), 24-25.

Sweller, J. (1990). On the limited evidence for the effectiveness of teaching general problem-solving strategies. *Journal for Research in Mathematics Education, 21*(5), 411-415.

Sweller, J. (1999). *Instructional design in technical areas. Australian Education Review* (Vol. 43). Camberwell, Vic: ACER Press.

Thompson, D., & Duncan, J. (1988). Using adventures in the classroom. *APCC Newsletter*(October).

Unwin, G. (1983). Adventures in education. *Creative Computing*(October), 149-157.

International Education Journal Vol 3, No 4, 2002
Educational Research Conference 2002 Special Issue
http://www.flinders.edu.au/education/iej

Developing Classroom Practices to Support NESB Students in Information Systems Courses: Some Preliminary Findings

Tracey Bretag
University of South Australia *Tracey.bretag@unisa.edu.au*

Sam Horrocks
University of South Australia *Sam.horrocks@unisa.edu.au*

Jeff Smith
University of South Australia *Jeff.smith@unisa.edu.au*

The number of international Non-English Speaking Background (NESB) students undertaking Information Systems (IS) courses at the University of South Australia has significantly increased in recent years. These students consistently achieve lower average grades than local students. This paper outlines the first cycle of an action research project which aims to facilitate improved learning outcomes for NESB students. The basic premise of the research is that international NESB students' academic performance is affected by socio-cultural and linguistic factors and that educational institutions have a responsibility to provide support to these students as they negotiate their new academic environment. Features of the project included training for IS tutors in English as a Second Language (ESL) teaching strategies, and a weekly support tutorial for students identified as needing assistance.

Information Systems, international students, Non-English Speaking Background (NESB), teaching strategies, action research

BACKGROUND

The number of international students (most of whom are of Non-English Speaking Background [NESB], and of Confucian Heritage Culture [CHC]) enrolled in University of South Australia (UniSA) programs of study has risen dramatically during the last decade, particularly in the Division of Business and Enterprise. Enrolment data shows that in two courses, Database Design (DD) and Business Systems Analysis (BSA), the number of international students has risen from 20 per cent of the student body in 2000 to 34 per cent in 2002.

Analysis of course results in Database Design and Business Systems Analysis show that as a group, international students are less likely to withdraw, but more likely to fail, and less likely to get a higher grade, than local students. DD results show a difference between local and international students' performance. Locals have a higher average mark and the gap between the two groups is getting wider (1.6 per cent in 2000, 5.3 per cent in 2001, and 5.8 per cent in 2002). BSA results show a difference between local and international students' performance, with locals again having a higher average mark (7.5 per cent, 13.5 per cent, and 8.4 per cent higher in the past three years).

The current research project was undertaken by Division of Business and Enterprise staff in two schools: Accounting and Information Systems, and International Business.

Sam Horrocks is the Program Director for Business Systems Analysis (BSA), and Jeff Smith is a lecturer in Database Design (DD); both teach in the School of Accounting and Information Systems. Tracey Bretag is a lecturer in Business Communication for Non-English Speaking Background (NESB) students in the School of International Business.

Despite perceived wisdom that Information Technology courses require less fluent English than other "language-rich" courses, Sam Horrocks intuitively felt that subjects such as Business Systems Analysis and Database Design do, in fact, require advanced levels of English. In addition to the difference in grades mentioned above, NESB students are often reluctant (or unable) to contribute to tutorial discussions. Furthermore, the written work submitted by NESB students for assessment is often difficult to read and mark, largely due to poor grammar.

Using Sam Horrocks' intuitive assumption (and the comparison of international and local students' results) as a basis, the research team asked the following questions:

- "What are the learning issues for NESB students in BSA and DD?"
- "What do linguists and other English as Second Language (ESL) educators have to say about these issues?"
- "Is there a particular methodology or methodologies which we might be able to adapt to meet the needs of our NESB students in Information Systems courses?"

The overarching question for the project was: "What specific teaching strategies could be used to assist NESB students to reach their academic potential in information systems courses?"

LITERATURE REVIEW

The research process began in late 2001 with a comprehensive literature review. We wanted to find out what other practitioners do/have done to assist NESB students undertaking business studies (specifically in information systems courses) in an Australian (or western) tertiary education institution.

Very little research has been done which specifically addresses the needs of students in information technology or information systems. Most of the literature focuses on students undertaking general business courses. The research tends to fall into three broad areas: 1) authors that argue that cultural difference is the primary factor influencing International (NESB/CHC) students' academic success in Western institutions; 2) those that suggest that language is just one factor, among many (including cultural and social isolation) that needs to be addressed; and 3) researchers that demonstrate that English language competence is the main factor affecting academic performance. All three positions maintain that there is a responsibility to provide wide-ranging institutional support to international (NESB) students to ensure that they reach their academic potential.

Cultural difference is the primary factor influencing international students' academic success.

Volet (1999) convincingly argues that, contrary to many educators' assumptions, Confucian Heritage Culture (CHC) students are not "spoon fed"; nor do they rely on a surface learning approach. She concludes that "The negative picture of Asian learners in Australian universities contrasts sharply with evidence from university statistics, that when English language proficiency is not an issue, Asian undergraduate students tend to perform better in their academic study than local students" (Volet 1999, p. 628).

The Chinese Learner: Cultural, Psychological and Contextual Influences (1996), edited by Watkins and Biggs, is a collection of articles which use a variety of methodologies to investigate the factors influencing CHC students in western learning environments. It is the variety of methods and perspectives offered by this book that makes it such an important contribution to the field, and which also challenge stereotypes of CHC learners as somehow "less able" than Australian students. A number of authors in this collection offer very practical classroom strategies to support CHC students, many of which deal specifically with developing English language competence:

1. Use an interactive, student-centred approach in tutorials (Tang 1996, p. 199)
2. Provide explicit expectations about assessment (Tang 1996, p. 199)
3. Provide the opportunity for collaborative learning, in particular for peer group discussion and peer tutoring (Tang 1996, p. 199 & Winter 1996, p. 221-122)
4. Provide opportunities for international students to work with Australian students (Biggs & Watkins 1996, p. 281)
5. Develop strategies to compensate for students' lack of English language fluency (Kirby, Woodhouse & Ma 1996, p. 155)
6. Facilitate students' access to the information in the text.
7. Elaborate discourse patterns, structures and rhetorical devices
8. Inculcate students into Australian academic culture (eg. Referencing, plagiarism, academic voice and register (Kirby, Woodhouse & Ma 1996, p. 143)
9. Provide assistance with discipline-specific vocabulary, and when necessary, Australian idioms, slang and cultural expressions (Kirby, Woodhouse & Ma 1996, p. 143).
10. Provide opportunities for students to use their English writing skills for tasks which are not for assessment (Tang & Biggs 1996, p. 170)

Weiland (1999) argues that "stress, language and general adjustment are not great problems. However, cultural differences do have a considerable impact" (p. 3). The author suggests that while students have the personal attributes that indicate that they will be successful academically (motivation, self evaluation, p. 2), the greatest impact on student learning is the new style of communication (in the Australian setting). Weiland concludes that academic skills preparation is essential to academic success for international students (p. 5).

Language is just one factor, among many to be addressed.

Batorowicz (1999), as part of a research project at the University of Southern Queensland, analysed the needs of two groups – Australian NESB and International students. The main issues that emerged were language difficulties, culture shock, homesickness and social difficulties. Batorowicz concluded that "…the process of internationalisation of Australian universities still faces many obstacles and…a formal policy on internationalisation or multiculturalism would benefit students from different cultures – and ourselves as well" (1999, p. 1).

Using survey data from 408 international undergraduate students and 121 staff, Robertson, Lane, Jones and Thomas (2000) concluded that international students experienced a range of challenges including social, administrative and linguistic.

Stoynoff (1996) conducted interviews with 27 international students at an American university and concluded that a strong support person is vital to the academic success of international students. Stoynoff's later (1997) research showed that "the vast majority of international students, even those with lower language proficiency, appear to succeed at

university. Therefore it is not as much a matter of whether international students will succeed but rather how successful they will be and at what personal cost" (p. 5).

Wicks (1996) compared the grades of 832 Australian students and 719 international students in both a mathematics and a language-rich course. The results of this research demonstrated that English competency is a not ultimately a problem for international students studying on campus in Australia, although offshore international students are significantly affected by lack of English language proficiency.

Using surveys, interviews and focus groups, Tompson and Tompson (1996) tabled the main issues faced by international students in two American business schools. These include: loneliness and fear; language; norms; rules and conventions; and cultural and academic environments.

Angelova and Riazantseva (1999) used participant and faculty interviews, observations, analysis of written samples and reflective journals kept by participants. Based on the various data, the authors concluded that international students need specific assistance to understand and adapt to the requirements of a western tertiary education environment.

Ballard and Clanchy (1997) provide some very useful advice in relation to facilitating students' learning in their book *Teaching international students. A brief guide for lecturers and supervisors*. These strategies include:

1. Establish names (including pronunciation) early in the semester (p. 38-39; see also Hellmundt, Rifkin & Fox 1998, p. 337)
2. Elicit responses, rather than just wait for a volunteer (p. 39)
3. Provide opportunities for success (Eg. Allow students time to discuss issues in pairs or small groups before speaking before the whole group).
4. Provide clear instructions for oral presentations (detailed information is provided on p. 40)
5. Use a "staged|" assessment schedule to enable students to build skills (p. 58).
6. Encourage students to take advantage of support services offered on campus.
7. Provide "model" answers that are easily accessible to all students (p. 62).

English language competence is the fundamental issue

Ruth Wajnryb, in her 2000 *Report of the English Language Services (ELS) Scoping Project* at the University of South Australia, consulted with a variety of staff, students and other stakeholders to provide an overview of the university's provision for NESB students. One key finding of the report is that an IELTS (International English Language Test System) score of 6.00 (the minimum standard accepted by the University of South Australia) is "barely adequate" for academic success.

Tracey Bretag uses the "Content-based ESL" in the development of her courses for NESB students in the School of International Business. This approach is based on collaborative curriculum development and teaching by a Language and Academic Skills (LAS) lecturer and a subject specialist. One advocate of the content-based model, Pantelides, argues in her 1999 article, "Meeting the needs of tertiary NESB students" that it is a university's responsibility to support students' language needs, and that an ESL specialist should teach in-faculty, where that faculty has a large proportion of NESB students.

Many researchers use practitioner research to provide a description and analysis of a course that they have developed and taught themselves (Allen & Rooney 1998; Beasley 1990 & 1997; Bretag & Scobie 2002; Chappel 1998; Chandrasegaran 1994; and Clerehan &

Crosling 1994). Without exception, each of the case studies concludes that English proficiency is the dominant issue faced by international students, in combination with socio-cultural factors.

Pearson (1999) and Beasley and Pearson (1999) have written numerous articles based on a longitudinal study from 1991-1997 at Murdoch University of two second-year business courses (with a high proportion of international students). The data included a comparison of average grades from 1992-1997, and also a comparison of grades for those who took an optional extra support tutorial and those who didn't. Results showed that students' grades improved significantly (from a 13 per cent failure rate in 1992 to a 1.5 per cent failure rate in 1997) due to a combination of improvements instigated by the authors. These changes included: an additional, voluntary learning support tutorial, team-taught by an LAS lecturer and subject specialist; a student centred approach; multiple, appropriate forms of assessment; timely and relevant feedback; and a focus on experiential activities.

Some specific strategies (in addition to a range of strategies similar to those by Ballard and Clanchy 1997, and Hellmundt, Rifkin & Fox 1998) suggested by Pearson and Bearsley (1999) are to integrate the teaching and developing of skills "in the context of the students' program of study rather than teach academic skills in isolation" (p. 308)

International students themselves have identified English proficiency as the key to academic success. Gatfield, Barker and Graham (1999) showed that of 26 variables, students are most concerned about academic instruction (in English).

Hellmundt, Rifkin and Fox (1998) provide practical classroom strategies based on research using an open-ended questionnaire, observation of students, informal and unstructured discussions and students' reflective journals. In addition to those suggested by Ballard and Clanchy 1997, Hellmundt, et al (1998) posits the following strategies:

1. Provide opportunities for students to speak as an "expert" (eg. About their own culture or personal experiences) (p. 336)
2. Allow students enough time to answer questions (p. 337)
3. Provide opportunities for international students to find out about Australian culture (p. 338).

From this perusal of the literature, it is clear that research in this field is premised on a number of key understandings. While some researchers focus on language competence, others are more concerned about cultural issues, and different learning backgrounds. All of the research, however, recognises that international (NESB) students require a range of institutional support arrangements; and that those responsible for teaching international students need to be cognisant of the myriad learning issues facing international students. Of significance to this project is the fact that there appears to be little or no literature addressing the particular needs of NESB students undertaking courses in information systems, or in courses generally perceived not to require high levels of English language competence.

Having looked at what other practitioners are doing to support their international (NESB) students, we decided to synthesise the key strategies, particularly those that were reported to have had a demonstrable impact on student learning outcomes.

METHODOLOGY

Dadds (1998, 41) suggests that:

> …practitioner research [refers] to forms of enquiry which people undertake in their own working contexts and, usually, on their professional work, in whatever sphere they practice. The main

> purpose of the enquiry is to shed light on aspects of that work with a view to bringing about some benevolent change.

Using Dadds' definition[1] as the basis for the project, the research team set out to instigate and/or adapt classroom practices and teaching strategies suggested by the literature, with the purpose of improving learning outcomes for NESB students in DD and BSA**.**

Fraser (1997, p. 169) argues "action research can be the most appropriate, most effective and least threatening strategy when evaluating curriculum innovations". Furthermore, Moller (1998, p. 71) suggests that genuine action research breaks down the binary between research and practice, and that useful action research is documented, published and scrutinised by peers. With this advice in mind, the research team has endeavoured to document the practice/research process, beginning with the first "cycle" of the project in Semester 1, 2002. As each "cycle" (semester) is completed, further reflection, documentation and peer review will take place.

Putting theory into practice

The strategic direction chosen by the research team is premised on the literature review, with specific attention given to a number of key texts, including Watkins and Biggs 1996 (including chapters by Tang; Kirby, Woodhouse & Ma; and Tang & Biggs), Hellmundt, Rifkin and Fox (1998) and Pearson and Beasley (1999). The following summarises the approach adopted in 2002:

1. Training in classroom teaching strategies for NESB students to lecturers and tutors in Business Systems Analysis and Database Design (the two courses under the direction of Sam Horrocks and Jeff Smith). This training took the form of a two-hour interactive workshop at the beginning of Semester 1, 2002. Teaching staff were asked to identify three strategies that they intended to use throughout the coming semester, and to be prepared to report back on if/how these strategies impacted on classroom participation
2. A team-taught support tutorial run on a weekly basis for students identified in the first tutorial as needing assistance with language and literacy. Students were asked to complete a short summary exercise of an IS-appropriate article.
3. Ongoing reflective journals to be kept by key participants (Sam, Jeff and Tracey) to be used as the basis for improving practise in the future, and for other research papers.
4. End of semester evaluation questionnaires to be distributed to participants of the support tutorial.
5. End of semester evaluation questionnaires to be distributed to tutors who had undergone the NESB teaching strategies workshop
6. Quantitative analysis in the form of a comparison between the final grades achieved by those who attended the support tutorials, and those who were invited to attend, but didn't.
7. A commitment to continue to provide tutor training and a weekly support tutorial for NESB students for at least four cycles (semesters) to ascertain the effectiveness of these strategies.

[1] Please note: In this paper, the terms "practitioner research" and "action research" are used interchangeably

THE TEAM-TAUGHT SUPPORT TUTORIAL

The tutorial is held once a week, with Sam Horrocks providing course content and Tracey Bretag providing "back-up support" in the form of language advice, information on how to go about assignments, prepare for examinations, etc. After a few weeks, the pattern of content settled into a rhythm of course content one week, followed by language and academic issues the following week.

The first time the support tutorial was offered (Semester 1, 2002), only a handful of students turned up (although 17 had been invited). This was due to a number of factors, including timetabling, mid-semester breaks, and a late start (week 3). We are now in the process of running the tutorials for a second time, and are currently adapting the process based on the lessons learned in the first cycle (eg. changes to the timing of tutorials and the introduction of "rewards" for attendance has increased attendance to 90 per cent of the students invited).

Qualitative Analysis: Student Evaluations

Who responded to the feedback survey?

All of the respondents had attended some of the workshops. Four of the respondents were regular attendees, one attended half the sessions and one responded who had dropped out of the workshops early in the semester. There were no responses from students who did not attend any of the support tutorials.

Without non-attendee feedback, the team had to develop on its own conclusions as to why the attendance was so low. The team concluded that students did not have enough incentive to attend. NESB students were already under pressure to complete the compulsory, assessable deliverables for their course and this was a purely voluntary tutorial.

Q. How did you feel when you received an invitation to attend the workshop?

In their first tutorial all students in the BSA course were asked to summarise a one-page article, which was then evaluated by Tracey Bretag. Those who were identified as having language and literacy difficulties (regardless of their linguistic background) were invited to attend the support tutorial.

The dilemma facing the team was that they would offend or upset students by using this process for selection; however, the team needed a mechanism to ensure that students were clearly identified as being in need of the workshop (note, Pearson and Beasley also used a summary exercise for early evaluation).

Bearing in mind that only those who attended responded, any inferences drawn form their response as to the impact of this process on the feelings of students cannot be generalised to the non-attendees. Interestingly, only one of the respondents indicated that they felt it was a reflection on them academically; the rest indicated that they were, in fact, expecting to be invited given their need to improve their language skills.

> I felt that perhaps more improvements in my English language have to be made.

Q. Did you find the workshops beneficial? In what ways?

Again, the lack of responses from non-attendees or those who attended once or twice makes it difficult to draw conclusions. All the respondents found the workshops beneficial. The first type of response was from those who felt the workshops provided an opportunity to "safely" explore concepts in more depth than a tutorial would provide.

> …the workshops were quite helpful, because it gave me an opportunity to ask questions and get further clarification on concepts I did not fully understand.

The other category of responses revolved around the study techniques aspects of the workshops.

> As I got lots of tips and guide not only related to the subject but also how to prepare them.

Q. What did you like about the workshops?

The majority of responses included a comment about the class atmosphere. The students felt that they were in a "safe" environment where they would not be mocked for asking "silly" questions or showing a lack of understanding of the concepts.

> I had less reservations about putting my hand up and saying I don't understand.

Another common and complementary theme was the size of the group. The students responded positively about the small size of the workshop and their subsequent increased ability to gain "attention" from the tutors ("*100 per cent attention*").

One student specifically addressed the issue of having two tutors in the room; one to provide support with language issues and one who was content focussed.

> I also liked how Tracey helped us to better understand what was being asked of us in a question, so that we were able to answer it correctly.

Only one student felt that the language support was not what they needed.

Q. What didn't you like about the workshops? How would you improve the workshops?

The responses varied enormously but some of the comments raised some interesting issues. Firstly, students would have preferred the support tutorials to be longer. The support tutorials were run for one hour per week, whereas the course tutorials were held every second week for two hours.

> I found the workshops to be not long enough, no sooner had we got into the class then we had to leave, I think they should be 2 hour workshops.

Some of the students expressed a desire to be provided with different questions in the support tutorial to allow them to further practice applying the concepts. This would provide them with additional feedback regarding their understanding of the material.

> I would prefer if it gives us some extra problems to make sure, the students grasp the concept of it…

Q. Do you feel the workshops had an impact on your ability to learn and succeed in the course?

The majority of students felt that the support tutorials had impacted positively on their ability to learn and succeed in the course in a positive way. Two of the students specifically addressed how it led to an improvement in how they approach tasks.

> Yes. I can now better understand what is required of me in the questions put to me by simply highlighting the keywords.

Any other comments

There appeared to be a genuine desire on behalf of the students to express their thanks at having the opportunity for additional support. Some respondents also indicated a desire for similar support workshops to be developed in other courses.

> The workshop was very useful and effective. Hopefully, there's going to be one for BSDI as well.

Qualitative Analysis: Tutor Evaluations

We had five casual tutors and two full-time staff respond to the survey. The discussion is divided into three main sections. Firstly, the evaluation of the tutor training workshop; secondly the teaching strategies themselves; and lastly, what tutors would do differently.

Q. Did you find the ESL teaching strategies training workshop useful?

The research team was interested in evaluating the usefulness of the training workshop held for tutors prior to the start of semester. The purpose of this workshop was to introduce the tutors to:

- the research being undertaken;
- the issues both staff and NESB students might face; and
- potential quick and easy strategies for dealing with some of those issues.

It was important that the strategies employed were not labour intensive for the tutor, did not intrude on class content time and would be well received by the students. The aim was to explore whether small changes in class teaching style would result in any benefit. The outcome of the tutor training workshop was for the two teams of tutors to devise a set of strategies they would implement across the course tutorials.

All of the tutors responded favourably to the workshop, with responses ranging from a straight "*yes they were useful*" response to a more detailed analysis of how they could be improved. Staff indicated that the session provided them with insight into the issues facing many NESB students as well as a planned approach for dealing with those issues or problems.

> Yes, I did not have prior knowledge or any real experience or appreciation of some of these issues. Good idea that we had a plan for the tutoring group and came up with ideas together.

Q. What ESL teaching strategies did you use in the class?; How do you believe the students responded to the strategies?; Did the strategies help you as a tutor?

Responses to these questions raised some interesting issues for the research team. What became clear was that although the tutors had agreed to implement three strategies, this was not necessarily what they did in class. Indeed, one tutor completely ignored the strategies agreed upon and took a completely different approach.

Most of the tutors employed at least two strategies, including the following:

- Learning names (name-tags or a name memory game).
- Subsequent use of names in class for asking and answering questions.
- Using small groups to check work within class, before whole class discussion.
- Structuring tutorials using an introduction, discussion and summary structure.

Only the first three strategies are discussed as the fourth strategy was not uniformly implemented.

Learning names and their subsequent use in tutorials

One of the easiest strategies involved making an attempt to learn the students' names. This was done either through taking digital photos of students, giving them name-tags or a "Name Memory Game". Six tutors utilised this strategy.

Results were mixed. Most tutors found the strategy worked well to increase participation of NESB students in class, provide the NESB students with more confidence in class and to make the class time more "enjoyable" for all students.

> I believe the students felt that the tutorials were more personal, less process driven.

Two tutors had difficulty with the implementation of this approach although they still found it a beneficial experience. A complicating factor was student volume. It was much harder to remember the students' names if there were more than 20 students in the class or if a tutor had a number of classes.

> I had 30 in one class and we could not even begin to play the game and there were too many students to remember who they were. It was held in a lecture theatre which made things worse as they spread out across the theatre.

The most successful approach was the use of digital photos taken in the first tutorial and then printed out and labelled with the student's name. This sheet was then either referred to before the class or during the class if required.

> I would get out the photos and look at them on the bus on the way in, this would help me with those I was having trouble remembering.

Regardless of the way in which they learned or identified names, most tutors found using students' names in class had a positive outcome, particularly in relation to class participation and "positive relationship development.

> Very positive. There was an increase of student participation voluntarily.

Using pair work and small groups to check work before whole class discussion

This was the other popular strategy. For most tutors it meant dividing the class into small groups to work on the answers before the wider class discussion ensued. There were multiple purposes to this strategy. Firstly, to give NESB students more confidence in their answers to questions; thus encouraging them to contribute to the wider class discussion. Secondly, to give NESB students an opportunity to discuss any queries or problems they might have in a smaller, less threatening forum. Finally, to allow NESB students to mix with local students with enhanced confidence and thereby develop relationships with other students in their course and program.

An increased contribution by NESB students was one of the key outcomes reported by tutors who used this strategy.

> … the main benefit was that they were in essence forced to contribute. And they saw that everyone was and I was not just asking them because they were quiet. It gave them more confidence I felt as they could double check their work with others…. Classes were better run as everyone was involved and I did not get the usual feeling that I was "pulling teeth".

One of the issues raised was that this could discourage students from preparing for class, and the implicit pressure on tutors to cover content.

> Many of the class and not just the NESB students had not done the work before hand (this was good and bad as I got the feeling by the end of semester that they then saw no need to do it as we would do it in groups anyway!!)… and I might add that it was very time consuming as the tutorials were two weeks worth [of content] in two hours so sometimes did not get through all the material.

Time also mentioned in a general context.

> but there was so much material to get through the students were more focused on the work than the Language exercises. I sensed a bit of frustration if I was off their tutorial material too much.

Overall comments

All the tutors indicated that they personally benefited from employing the strategies in class.

> Yes [they helped me as a tutor] – but more the enlightenment that; different cultures respond to lecturers (authority figures) in different manners. I believe the subject that we facilitate requires experience and understanding, it requires issue interpretation rather than textbook knowledge.

> The limited class involvement from some international students and the acceptance of the tutors word as gospel, I believe limits the students ability to experience the subject, particularly when the students interpretation may be a viable alternative which could be explored in class.

What would you do differently next time?

Most tutors commented that the strategies were easy to implement and indicated no specific changes they would make.

Quantitative analysis

We investigated the results by dividing the *Business Systems Analysis* class into three groups:

- Identified students that attended (IA);
- Identified students that didn't attend (InA);
- Others (O).

BSA S1 results show a consistent pattern in the results for these groups with *Others* (generally Australian, local students who were not identified as requiring support) achieving an overall higher result than the students who were invited to attend (generally international, NESB students). Of those invited to attend the support tutorial, those who did attend achieved an overall higher result than those who didn't ($Mark_O > Mark_{IA} > Mark_{InA}$). The overall average mark in the course for the three groups was 64 per cent, 60 per cent and 52 per cent respectively.

Analysing the results

The research team is reluctant to attribute these grades specifically to the strategies employed in just one semester; particularly given the small number of students who attended the support tutorial, and the resultant lack of data. However, the preliminary findings are positive and support the need for further research. It should be noted that the research undertaken by Beasley and Pearson (1999), the results of which provided the basis for this project, was a progressive study over a seven-year period. It would be unreasonable to expect any significant changes to student learning outcomes in just one semester.

To put these results in context, the research team decided to look at these students' entire academic history at UniSA. The average result of all courses studied in their programs by the two groups of identified students indicates that those who *did* attend the support tutorial, achieved a higher average mark (64.9%) than those who were invited but didn't attend (54.3%).

At this stage it is difficult to conclude much from the data, except perhaps to suggest that NESB students who are pro-active about their learning and who take advantage of every support opportunity, may be more likely to succeed than those who do not access support services. However the students' different academic histories (types of courses undertaken, length of study, etc) makes any conclusion purely speculative.

CONCLUSION: IMPLICATIONS FOR THE NEXT RESEARCH CYCLE

Staffing

Commitment to the research is difficult with contract staff. The increasing use of contract staff:

- makes the training of tutors in use of strategies resource intensive;

- does not allow for control over the actual implementation of the strategy, thus making it impossible to ensure conformity across courses;
- does not encourage commitment to this kind of research from tutors; they are paid a relatively small amount of money to do an ever-increasingly time intensive task and this is a further burden on their time; and
- makes the retrieval of feedback difficult (for the reasons given above).

ESL Strategies

Even small changes can have a positive impact. The learning environment has been particularly influenced with NESB students participating more and feeling included and respected. Many tutors may be unaware of the issues and hence unaware of particular strategies they might employ – we should establish tutor training in this area for a wider range of courses.

The support tutorial

Based on the analysis of available data, the research team has concluded that students need an incentive to attend support tutorials, and the small size of the tutorials is a highly beneficial aspect of their design. The results of changes made in the next cycle of activities will address these assumptions and provide more data for analysis.

Resource issues are paramount to the future success of the support tutorials. The team-teaching approach, integrating language/academic skills and content is a well-researched model which students have commented favourably upon. Furthermore, students would like the tutorials to be longer and to be provided with additional material. If we continue the support tutorials, increase their length and prepare additional materials, we will clearly need a substantial source of funding. For this to occur, there needs to be a commitment at all levels of the institution.

REFERENCES

Allen, R., & Rooney, P. (1998). Designing a problem-based learning environment for ESL students in business communication. *Business Communication Weekly, 61*(2), 48-56.

Angelova, M., & Riazantseva, A. (1999). 'If you don't tell me, how can I know?' *Written Communication, 16*(4), 491 (435p).

Ballard, B., & Clanchy, J. (1997). *Teaching International Students: A Brief Guide for Lecturers and Supervisors*. Deakin, ACT: IDP Australia.

Batorowicz, K. (1999). A fair go? The problems and needs of international and non-English speaking background students at Australian universities. *Youth Studies Australia, 18*(3), 37 (34p).

Beasley, C. J. (1990). Content-based language instruction: Helping ESL/EFL students with language and study skills at tertiary level. *TESOL in Context, 1*, 10-14.

Beasley, C. J. (1997). *Students as teachers: The benefits of peer tutoring*. Teaching and Learning Forum.

Beasley, C. J., & Pearson, C. A. (1999). Facilitating the learning of transitional students: Strategies for success for all students. *Higher Education Research & Development, 18*(3), 303-321.

Biggs, J. B., & Watkins, D. A. (1996). The Chinese learner in perspective. In D. A. Watkings & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 269-281). Hong Kong and Melbourne: CERC & ACER.

Bookin-Weiner, J. B. (1995). Summary of Group Two: Academic issues. In H. de Wit (Ed.), *Strategies for Internationalisation of Higher Education* (pp. 92-93). Amsterdam: EAIE.

Bretag, T., & Scobie, S. (2002). *Innovative classroom practice in an offshore environment.* Paper presented at the HERDSA Quality Conversations, Perth, Western Australia, 7-10 July.

Chandrasegaran, A. (1994, 21-22 November). *Integrating content-course tasks into the teaching of writing skills for academic purposes.* Paper presented at the Integrating the teaching of academic discourse into courses in the discipline, Melbourne.

Chapple, S. (1998). *How do you foster enthusiasm and participation in tutorials particularly by international students?* Teaching and Learning Forum.

Clerehan, R., & Crosling, G. (1994, 21-22 November). *A case study of an adjunct program: English for first-year engineering students.* Paper presented at the Integrating the Teaching of Academic Discourse into Courses in the Disciplines, Melbourne.

Gatfield, T., Barker, M., & Graham, P. (1999). Measuring student quality variables and the implications for management practices in higher education institutions: An Australian and international student perspective. *Journal of Higher Education Policy & Management, 21*(2), 239 (216p).

Hellmundt, S., Rifkin, W., & Fox, C. (1998). Enhancing intercultural communication among business communication students. *Higher Education Research & Development, 3*(3), 333-344.

Kirby, J. R., Woodhouse, R. A., & Ma, Y. (1996). Studying in a second language: The experiences of Chinese students in Canada. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 141-158). Hong Kong and Melbourne: CERC & ACER.

Leask, B. (2001). *Teaching NESB and international students at the University of South Australia*. University of South Australia [2001, August].

Pantelides, U. (1999). Meeting the language needs of tertiary NESB students. *Australian Journal of Language and Literacy, 22*(1), 60-75.

Pearson, C. A. (1999). *Workshops as stimulating learning environments.* Teaching and Learning Forum [2001, 17 December].

Robertson, M., Lane, M., Jones, S., & Thomas, S. (2000). International students, learning environments and perceptions: A case study using the Delphi technique. *Higher Education Research & Development, 19*(1), 90-102.

Scarino, A. (2001). *Internationalisation of the curriculum - structural pathways within and between courses.* University of South Australia [2001, August].

Stoynoff, S. (1996). Self-regulated learning strategies of international students: A study of high and low achievers. *College Student Journal, 30*(3), 329 (328p).

Stoynoff, S. (1997). Factors associated with international students' academic achievement. *Journal of Instructional Psychology, 24*(1), 56 (13p).

Tang, C. (1996). Collaborative learning: The latent dimension in Chinese students' learning. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 197-203). Hong Kong and Melbourne: CERC & ACER.

Tang, C., & Biggs, J. B. (1996). How Hong Kong students cope with assessment. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 159-182). Hong Kong and Melbourne: CERC & ACER.

Tompson, H., & Tompson, G. (1996). Confronting diversity issues in the classroom with strategies to improve satisfaction and retention of international students. *Journal of Education for Business, 72*(1), 53 (55p).

Volet, S. (1999). Learning across cultures: Appropriateness of knowledge transfer. *International Journal of Educational Research, 31*(7), 625-643.

Wajnryb, R. (2000). *Report of the English Language Services (ELS) Scoping Project.* University of South Australia, Division of Education, Arts & Social Sciences.

Watkins, D. A., & Biggs, J. B. (1996). *The Chinese Learner: Cultural, Psychological and Contextual Influences.* Melbourne: CERC & ACER.

Weiland, R., & Nowak, R. (1999). *Academic preparation programs: A schema approach to learning in context*. Teaching and Learning Forum.

Wicks, R. (1996). Effects of English language proficiency on the academic performance of international students: A US study. *Distance Education, 17*(1), 196-204.

Winter, S. (1996). Peer tutoring and learning outcomes. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 221-241). Hong Kong and Melbourne: CERC & ACER.

The Impact of IELTS Scores on Performance at University

Vicki Feast
University of South Australia *vicki.feast@unisa.edu.au*

This paper considered two key issues. First, the relationship between English language proficiency, as measured by IELTS tests scores, and performance at University, as measured by Grade Point Average (GPA), was investigated using multi-level analysis. A significant and positive relationship was found. Second, the trade-off between raising IELTS entry scores and the consequent loss of international students was investigated at one South Australian university. Recommendations for raising IELTS scores for undergraduate and postgraduate students were made in the paper although it was recognised that on financial grounds the loss of international students might be too large to justify the increase in minimum English proficiency standards.

English proficiency, IELTS, academic performance, GPA, multilevel analysis

INTRODUCTION AND SIGNIFICANCE OF THE TOPIC

The proportion of funds that derive from government sources to support the core business of teaching students in universities is declining on a per capita basis. "Since the mid 1980s the higher education sector has grown at a rate faster than the level of Commonwealth funding for the sector" (Nelson, 2002, p.17) and "A return to full public funding of Australian universities will not occur. This would require a further $4 billion annually of Commonwealth funding." (Nelson, 2002, p.v). Therefore, income from other sources including "income from fee-paying students is an important and growing source of university revenue[1]" (Australian Vice Chancellors' Committee (AVCC) Fact Sheet 2, 2001, p.1). It follows that universities are increasingly dependent on alternative non-government sources of funding and full fee paying international students are a large and increasing source of that revenue[2]. Consequently, the financial future of Australian universities may well depend on the trend to enrol ever increasing numbers of international students who study both in Australia and offshore.

Additional to this funding imperative, there is both a moral and financial need to ensure that fee-paying international students who undertake university education are capable of succeeding. Hence, as for all university studies, standards are required to enable international fee-paying students to gain entry into university. These standards include academic entrance levels which are derived from a students' previous studies, work experience and academic entrance tests. In the case of international students from non-English speaking backgrounds, there is also a requirement that they have a minimum

[1] In 1992 Fee-paying students contributed 5 per cent to university funding sources and in 1999 this rose to 12 per cent (AVCC Fact Sheet 2, 2001, p.1)
[2] Overseas fee paying students contribute 79 per cent of university fee-paying revenue (AVCC Fact Sheet 2, 2001, p.1)

standard of English proficiency for entrance into Australian courses and programs that are taught in English.

One of the most popular and most used tests of the standard of English is the International English Language Testing System (IELTS). IELTS scores are required for some students from particular countries[3] to gain their visas to enter Australia and Australian universities typically require students to have IELTS scores (or equivalent English proficiency standards) to gain entry. Although there are other English language proficiency tests available and in use, IELTS is the best regarded by Australian universities and the Australian government. The minimum IELTS score for university entrance is typically an average score of 6.0 (or equivalent) although some universities require higher scores for particular programs and postgraduate studies. "In general an Overall Band Score[4] between 6.0 and 7.0 in the Academic modules is accepted as evidence of English language proficiency for higher education institutions around the world" (Ciccarelli, 2001, p.1).

Despite these requirements there is a degree of disquiet amongst university colleagues about the English language abilities of some international students. In staff rooms and meetings they can be heard to be complaining about marking scripts of students whose English is below the standard that they consider acceptable for university study. Indeed, the recommendations published by IELTS Australia state that an IELTS score of 6.0 requires further English study for linguistically demanding academic courses and also linguistically less demanding academic courses and is said to be "acceptable" only for linguistically less demanding training courses. In fact, the IELTS guidelines recommend an IELTS score of 7.0 as "probably acceptable" for linguistically demanding academic courses and "acceptable" for linguistically less demanding courses (IELTS, 2001, p.22). It is clear that significant numbers of students are being admitted to Australian universities at a level below that suggested as acceptable by IELTS Australia.

This apparent contradiction between the levels of English deemed acceptable by most universities and the IELTS recommendations, and indeed the view of a growing number of academics, warrants further investigation. There is also a growing body of literature that suggests a link between English language proficiency and academic success:

> In relation to the IELTS test a number of predictive validity studies have been carried out which conclude that language proficiency is a critical factor in academic success and the IELTS is a useful predictor of a student's ability to cope with academic English (Ciccarelli, 2001, p.3).

Whilst the use of the word critical in the Cicarelli quote may overstate the statistical significance of the link between language proficiency and performance at university, it is indicative of the importance with which English language proficiency is viewed in some Australian universities.

However, Graham (1987, p.517), Burns (1991, p.75) and Dooey (1999, p.115) recommend that each institution conduct its own studies concerning the link between English proficiency levels and academic success and make its own decisions about acceptable English language proficiency levels. This suggestion is made for a number of reasons including the difficulty of generalising findings from previous studies and the limitations of these studies, many of

---

[3] Recent changes in legislation require students from Categories 3 and 4 countries (such as Argentina and China) to have a minimum average IELTS score of 6.0 (or 5.0 IELTS and 30 weeks of English language study) to obtain a visa for tertiary study. Since July 1st, 2001 TOEFL or other equivalent tests are no longer acceptable for student visa purposes, unless gazetted by Parliament, as IELTS tests are viewed by the Department of Immigration as the test with the most integrity.

[4] The Overall Band score is the average of all four modules or sub-test items. Thus it is like an average test score for IELTS.

which are small-scale. In addition, it is argued, since there were a number of changes to the IELTS test in 1995 and much research was conducted prior to 1995, consequently further research is needed. The present study contributes to the field of research on English proficiency standards by using new statistical techniques to examine further the relationship between IELTS scores and academic performance for international students admitted in 2000 and 2001 to one university.

It is widely understood by academics and researchers that the reasons why students perform poorly at university are many, varied and complex. Personal background factors (for example age, gender, personality, attitude, and motivation), academic background factors and requirements (for example, previous studies within certain fields), teaching and support factors (for example, English language support and general study advice, the interest and ability of teaching staff) cultural factors and language proficiency factors all have a role to play in influencing how successful students will be in their tertiary study. For obvious reasons, particularly for international students cultural and language factors are prominent in the literature. Hill, Storch and Lynch (1999, p.63) claim "Nobody would argue that ELP [English Language proficiency] has no role to play in academic achievement". This study focuses on the impact of English language proficiency on performance.

LITERATURE REVIEW

This study focused on linguistic factors affecting performance by investigating the impact of English language proficiency tests on academic performance. Previous research has produced varied results. A number of studies (Criper and Davies, 1988; Elder, 1993; Ferguson and White, 1993; Cotton and Conrow, 1998; Hill et al, 1999; Kerstjens and Nery, 2000) have found a weak positive association between IELTS and GPA. A typical standardised regression coefficient of around 0.3 was common among many of these studies. However, in an Australian context at the University of Melbourne, Hill et al (1999, p.55) stated that that "the relationship between Grade Average and IELTS score was found to be moderately strong (r = 0.540)". To put this result in context, Hill et al (1999, p.55) also concluded "An examination of the various scatterplots suggested a violation of certain assumptions of the regression model".

At the University of Tasmania, Cotton et al (1998, p.98) found "a weak (positive) association for the Reading and Writing subtests (modules)" but "very low or negative associations between academic performance and the other subtest scores, as well as between IELTS global (average) scores and academic performance overall". Similarly, Kerstjens et al (2000, p.105), using RMIT student data found a "small-to medium predictive effect of academic performance from the IELTS score for the … Higher education group, accounting for ... 9.1 per cent of the variation in academic performance". At the University of Hong Kong, Ho and Spinks (1985, p.249) found that "English language skills had the most predictive value [compared to various factors such as intelligence and personality variables], accounting for about 10 per cent of the variance of performance measures.

Some studies found no statistically significant relationship between IELTS and academic performance (Traynor, 1985; Fiocco, 1987; Graham, 1987; Light, Xu, & Mossop, 1987; Gibson and Rusek, 1992; Rusek, 1992). In addition, others found their results inconclusive (Dooey, 1999, p.114). Dooey (1999, p.117) found that there was no evidence to suggest that students who did not meet the entry criteria (IELTS 6.0) were destined to fail but conversely most of the failures were students who entered Curtin University with high IELTS scores.

Moreover, it proved more difficult to find recent studies that predicted strong associations between the two variables. Dooey (1999, p. 115) claimed "most [studies] did not find a

strong overall positive association between IELTS scores and subsequent academic success". The exception seemed to be a study by Bellingham (1993, p.229) conducted in New Zealand that revealed a moderate association partly because of its unique nature in that it included students with a wide range of IELTS scores including some below 5.0. Most other studies were unable to include students with IELTS scores below 6.0 as this was the minimum entry standard required for university admission. The Bellingham study (1993, p. 231) found that "If these results were generalisable to the wider population, this would mean that with scores below 6.0 in IELTS, students have a 20 per cent chance of passing; whereas at 6.0 or more the chance is 50 per cent". Bellingham went on say that this finding was consistent with IELTS guidelines that at 6.5 "language proficiency is less likely to be a significant factor in influencing academic success". However further discussion in the Bellingham paper confirmed that factors other than language are also "integral to academic success".

Gibson and Rusek (1992, p.17) suggested that the contradictory results of various studies did not indicate that the tests were not valid measures of English proficiency but that "language skill is only one of the variables which predicts academic success, albeit an important one". They recommended that each institution should carry out its own research into acceptable English language levels based upon past student achievements.

In relation to the second research question, a minimum English language proficiency score of 6.0 IELTS has been established at a number of Australian universities. This level was informed by research and levels set by competing tertiary institutions. Ferguson et al from Edinburgh University have stated (1993, p.34) "Band 6 seems to represent some sort of cross-over line". Criper and Davies (1988, p.79) were also reported in Ferguson et al (1993, p.35) to have stated that Band 6 is "some kind of changeover score". Ferguson et al (1993, p.36) concluded "there is a level of proficiency below which failure increases sharply, and that language assumes a more important role in academic performance when proficiency is low." The study by Elder (1993, p.87) contradicted these findings to some extent as it indicated that "the strongest level of agreement between test predictions and academic outcomes occurred at Band 4.5 [which] casts some doubt on the recommendation that a Band score of 6.5 [should] be a minimum requirement".

RESEARCH QUESTIONS

This paper investigates two research questions. The first research question is What is the relationship between English language proficiency levels (as measured by IELTS scores) of international students and their academic performance at University, as measured by GPA? This question considers the degree to which IELTS test scores are meaningful predictors of success at university in 2000 and 2001.

The second research question considers whether the current minimum entrance score of IELTS 6.0 should be increased. The purpose of this research question is to examine whether there is a need to raise the minimum IELTS score that international students should have in order to gain admittance to university courses so they have a reasonable chance of success. This research question is investigated for both postgraduate and undergraduate students. The links between the two research questions is also investigated later in this paper.

METHOD

The method used to undertake the first research question in this study involved a commonly used technique called regression analysis.

> Regression analysis refers to a broad class of statistical techniques that are designed to study the relationship between a criterion (or dependent variable), *Y*, and one or more predictors (or independent variables), $X_1, X_2,\ldots,X_p$. (Tatsuoka, 1997, p.648).

In this study the dependent variable was the performance of international students at university as measured by their Grade Point Average (GPA). The key independent variable was English proficiency as measured by students' IELTS scores on entry to university. The proposition addressed in this study was that there is a relationship between English proficiency levels (overall IELTS) and performance at university (GPA). Indeed, it was postulated that IELTS is positively related to performance of international students at this university.

Multiple regression analysis was undertaken in this investigation with the dependent variable of mean GPA and with the independent variables: IETLS and age (a continuous variable), semester of entry, division (broad discipline area of study), home country, gender, and level of study (postgraduate and undergraduate) which were categorical variables. Thus the general equation for this analysis was

GPA = f [5] (English proficiency, age, gender, entry semester, discipline area, home country, level of study)

A particular type of regression analysis called multilevel analysis was conducted using GPAs gained progressively through Semester 1 to Semester 5. Five semesters were the maximum number of semesters studied by the international students who were accepted into the university in 2000 and 2001. Some students had only one mean GPA statistic for one semester if they enrolled in Semester 2, 2001 and others had up to five mean GPA scores if they enrolled in Semester 1, 2001 (one for each semester of enrolment in the two years and including a summer school). The multi level analysis used in this research was based upon models where sampling and measurement errors were estimated in a hierarchical manner at two levels. Use of this class of models assisted in the estimation of the impact of English language proficiency, as measured by IELTS test scores, over time as students progressed through their studies, while controlling for the effects of age, gender, entry semester, discipline area, home country and level of study. Multilevel analysis permitted a more appropriate and detailed intra and inter-student analysis of the relationship between IELTS and GPA than is possible with simple regression analysis using mean GPA scores.

The purpose of the multilevel analysis was to investigate the relationships between the variables, English language proficiency scores (IELTS) and GPA, at the intra and inter-student levels, after controlling for other factors that might influence GPA scores. The multilevel analysis was conducted using the hierarchical linear modelling (HLM) procedures (Bryk and Raudenbush, 1992). HLM made it possible to analyse variables at intra and inter-student levels simultaneously so that the impact of these variables on GPA were examined in one analysis. A further strength of HLM was that it was also possible to investigate the interaction effects between the variables at the two different levels.

According to Hox (1994, II) "multilevel regression models are essentially a multilevel version of the familiar regression model". Hox (1994, p.5) also stated "A multilevel problem is a problem that concerns the relationships between variables that are measured at a number of different hierarchical levels". Table 1 lists and explains the variables used in the analysis.

[5] Where f is an unknown function but assumed to be linear for the purposes of the analysis

Table 1. Explanation of variables used in analysis

Variable	Score range and values	Notes
CRITERION		
GPA	1.43.-7.0	Mean Grade Point Average for students for a semester of study.
LEVEL 1		
OCCASION	Up to five time points	Changes to GPA over time.
LEVEL 2		
IELTS	4.5 – 8.5	International students' (recorded on the ISIS database) IELTS mean overall test score in either 2000 or 2001. These scores are the mean of four individual module scores for Reading, Writing, Speaking and Listening.
IELTS module scores	4-8	Individual test scores for each of the modules (Reading, Writing, Speaking and Listening). Their average is the overall mean IELTS score.
AGE	17-40	Age of student (in years).
GENDER	0 = male, 1 = female	Sex of student (Male or Female).
COUNTRY	1-35	Home country of student.
CHINA	0 = all other countries 1 = born in China	Students born in China as a variable.
VIET	0 = all other countries 1 = Vietnamese born	Students born in Vietnam as a variable.
LEVEL	0 = undergraduate 1 = postgraduate	Level of study - Postgraduate or undergraduate study.
DISCIPLINE	BUE, EAS, ALS, IEE, HSC	Broad area of study divided into six disciplines as defined by faculty (division) structures.
BUE	0- = all other disciplines 1- = students from BUE (Business and Enterprise Division)	Students from the Business and Enterprise Division as a variable.
EAS	0 = all other disciplines 1 = students from EAS (Education, Arts and Social Science Division)	Students from the Education, Arts and Social Science Division as a variable.
HSC	0- = all other disciplines 1 = students from HSC (Health Science Division)	Students from the Health Sciences Division as a variable.
IEE	0 = all other disciplines 1 = students from IEE (Technology and Engineering Division)	Students from the Technology and Engineering Division as a variable.
ALS		Students from the Access and Learning Support Unit.

The student group included in this study were 101 international students, from a total of 964 international onshore students, who were admitted to one university on the basis of their IELTS score during 2000 and 2001. The IELTS student group characteristics can be summarised as follows:

- approximately 50 per cent were males and 50 per cent were females

- approximately 50 per cent were postgraduate and 50 per cent were undergraduate students.
- they were aged between 17 and 40, with a mean age of 25.6 years.
- they originated from 30 different countries, with Malaysian, Indonesian, Thai and Chinese students dominant.
- they were studying in five disciplines areas with most students (47%) enrolled in a business faculty.
- they had GPA scores ranging from 1.5 to 7.00, with a mean GPA of 4.87.
- they had IELTS scores ranging from 4.5 to 8.5, largely between 6.0 and 7.0.

The second research question was investigated in this study by conducting an analysis of the trade-off between the loss of student numbers experienced by raising the current minimum scores (from IELTS 6.0) and the subsequent increase in GPA of the remaining students. Various methods of selection by raising the scores were tested including raising the overall scores and imposing some minimum requirements for the module scores. Examples included raising the IELTS overall score to 6.5 for undergraduates and 7.0 for postgraduates. Five different methods of selection were tested for undergraduate students and six alternative methods of selection were tested for postgraduate students. The loss of students, calculated as the percentage of students who would not have gained entry to the university using the higher scores for each method of selection, was considered against a calculation of the rate of increase of GPA. The rate of GPA increase was calculated by comparing the GPA of all students who entered the university in 2000/2001 with an IELTS score with the GPA of the smaller group who would have been admitted under the higher score requirements. As the appropriate minimum English proficiency score was largely a political decision to be made by university management, this trade-off was considered an appropriate way to analyse this question.

CONTEXT

IELTS was “designed to assess the language ability of candidates who need to study or work where English is used as the language of communication” (IELTS, 2001, p.1). The test was originally designed at the University of Cambridge and jointly developed with the British Council in 1980 but now can be taken in “251 test centres in over 105 countries” (IELTS, 2001, p.4). It has been widely accepted and used in Australia since 1989. In 1995, a number of significant changes were introduced to address concerns with practical issues, administrative problems, technological developments and theoretical issues (Charge and Taylor, 1997, p.379).

At the present time, an IELTS score of between 6.0 and 7.0 is commonly accepted at Australian universities as the minimum score necessary to demonstrate English language proficiency for admitting international students. The IELTS has two components: Academic or General with four modules: Reading, Writing, Listening and Speaking. The Speaking and Listening modules are common to both the Academic and General components, but the Reading and Writing modules are different for each component. The Academic module is used for entry into universities, as the focus of the General module is “on basic survival skills in a broad social and economic context” (IELTS, 2001, p.4). Students who sit the IELTS are scored in Bands from zero (did not attempt the test) to nine (expert user).

MULTILEVEL ANALYSIS FINDINGS

The models presented below for IELTS students are the null model (Table 2) and the final and best fitting model (Table 3). The multi-level analysis undertaken in this study is an analysis of the change data at two distinct Levels. At the first Level, the regression relationship examines how student performance (GPA) changes over time. At the second Level, the factors influencing average student performance (GPA) and change in GPA over time are examined. Multi-level analysis is used as it produces a model that controls for change in GPA over time and also controls for factors that may affect change in GPA over time. Other studies, the results of which are outlined earlier in the Literature Review and which use less effective regression analysis, are not able to control statistically for those factors which confound the relationship between IELTS scores and GPA. Using multi-level analysis, and with this control of factors, the relationship between IELTS scores and average GPA is made clearer.

Table 2. Null model for IELTS students

Summary of the model specified (in equation format)

Level-1 Model Y = B0 + R

Level-2 Model B0 = G00 + U0

The outcome variable is GPA

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	4.88	0.11	44.78	84	0.000

Deviance = 655.39293

Number of estimated parameters = 1

Table 3 shows no significant effects at Level 1 (the micro level within the 97 students in the sample) but some significant effects at Level 2 (the macro level between students). At Level 1, the variable OCCASION (B_1) has virtually no effect on GPA for most students. However, both the variables IEE (students from the Information Technology and Engineering discipline) and VIET (students born in Vietnam) have a significant interaction with OCCASION (at the $p < 0.05$ level of statistical significance). Table 3 shows that at Level 2, IEE has a significant interaction (at the 5% level, p-value of 0.030, t ratio of 2.17) with OCCASION (changes in GPA over time during the five semesters in 2000 and 2001) given that the model described in Table 3 is the best fitting model for IELTS students. With IEE as one of two predictors and given a positive regression coefficient of +0.27, this indicates that the 25 engineering/technology students with an IELTS score have a GPA that is rising over time when compared with other students. This also shows that IEE students are improving on their GPA over time because either their language proficiency levels are becoming less important as they proceed through their program or that having the opportunity to study in Australia leads to improving language proficiency skills. This interaction effect is displayed in Figure 1 below.

Table 3 also shows that at Level 2, the variable VIET (7 students born in Vietnam) has a significant interaction (at the 5% level, p-value of 0.000, t ratio of –3.74) with OCCASION (changes in GPA over time during the five semesters in 2000 and 2001) given that the model described in Table 3 is the best fitting model for IELTS students. With VIET as one of two predictors and given a negative regression coefficient of –0.58, this indicates that Vietnamese born students with an IELTS score have a GPA that is declining over time when compared with all other students. This is of concern as it shows that VIET students have a

worsening GPA over time because either their language proficiency levels are becoming more important as they proceed through their program or studying in Australia has led to a deterioration of their language proficiency skills. This interaction effect is also displayed in Figure 1.

Table 3. Final model for IELTS students

Level-1 Model Y = B0 + B1*(OCC) + R

Level-2 Model B0 = G00 + G01*(IELTS) + G02*(LEVEL) + G03*(CHINA) + G04*(HSC) + U0

B1 = G10 + G11*(IEE) + G12*(VIET) + U1

The outcome variable is GPA

Final estimation of fixed effects:

Fixed Effect	Standard Coefficient	Approx. Error	T-ratio	d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	4.30	0.14	31.69	92	0.000
IELTS, G01	0.39	0.13	2.92	92	0.004
LEVEL, G02	0.79	0.17	4.52	92	0.000
CHINA, G03	0.99	0.41	2.41	92	0.016
HSC, G04	1.20	0.35	3.48	92	0.001
For OCC slope, B1					
INTRCPT2, G10	0.01	0.07	0.20	94	0.843
IEE, G11	0.27	0.12	2.17	94	0.030
VIET, G12	-0.58	0.15	-3.74	94	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	0.75	0.56	64	374.47	0.000
OCC slope, U1	0.19	0.04	66	72.58	0.270
level-1, R	0.62	0.38			

Statistics for current covariance components model

Deviance = 566.5804

Number of estimated parameters = 4

Figure 1 shows the combination of these two effects (VIET and IEE).

At Level 2 (B_0), all of the dependent variables shown in Table 3 are significant (at the $p < 0.05$ level). The variable IELTS, with a p-value of 0.004 and a t ratio of 2.92, has a significant relationship with GPA. A regression coefficient of +0.39 for IELTS indicates that there is a positive relationship between the IELTS score and GPA. Thus higher IELTS scores are related to higher mean GPA scores. For every one-unit increase in IELTS scores, assuming other variables are kept constant, mean GPA increases by 0.39. The regression coefficient of +0.39 (and a t-ratio of 2.92) indicate that the relationship is relatively weak but in line with other studies (Criper and Davies, 1988, Ferguson and White, 1993, Elder, 1993, Cotton and Conrow, 1998, Hill et al, 1999, Kerstjens and Nery, 2000).

The variable, LEVEL (level of study), has a p-value of 0.000 and a regression coefficient of +0.79, and with a t-ratio of 4.52, indicates a strong positive relationship between level of study and mean GPA for IELTS students. In practice it can be argued that IELTS students with a higher level of study (postgraduate students) are associated with higher mean GPA

scores. Postgraduate students on average have a higher mean GPA (by 0.79) compared to undergraduate students after adjustments are made for the other variables. Thus postgraduate students with IELTS scores have a greater chance of success than undergraduate students with the same IELTS scores, all other variables being equal.

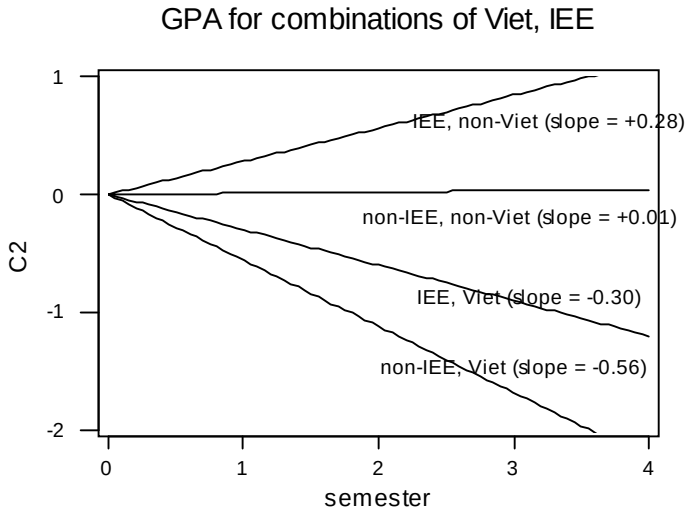


Figure 1: Interaction effect for IELTS students

The variable, CHINA (7 students whose home country is China), has a p-value of 0.016 and a regression coefficient of +0.99 (with a t-ratio of 2.41) indicating a relatively weak relationship between students from China (with IELTS scores) and GPA. This indicates that on average students from China have a higher GPA (by 0.99) than students from other non-Chinese countries after adjustments are made for other variables. In practice this shows that Chinese born students are likely to be relatively more successful than non-Chinese born students with the same IELTS score, all other variables being equal.

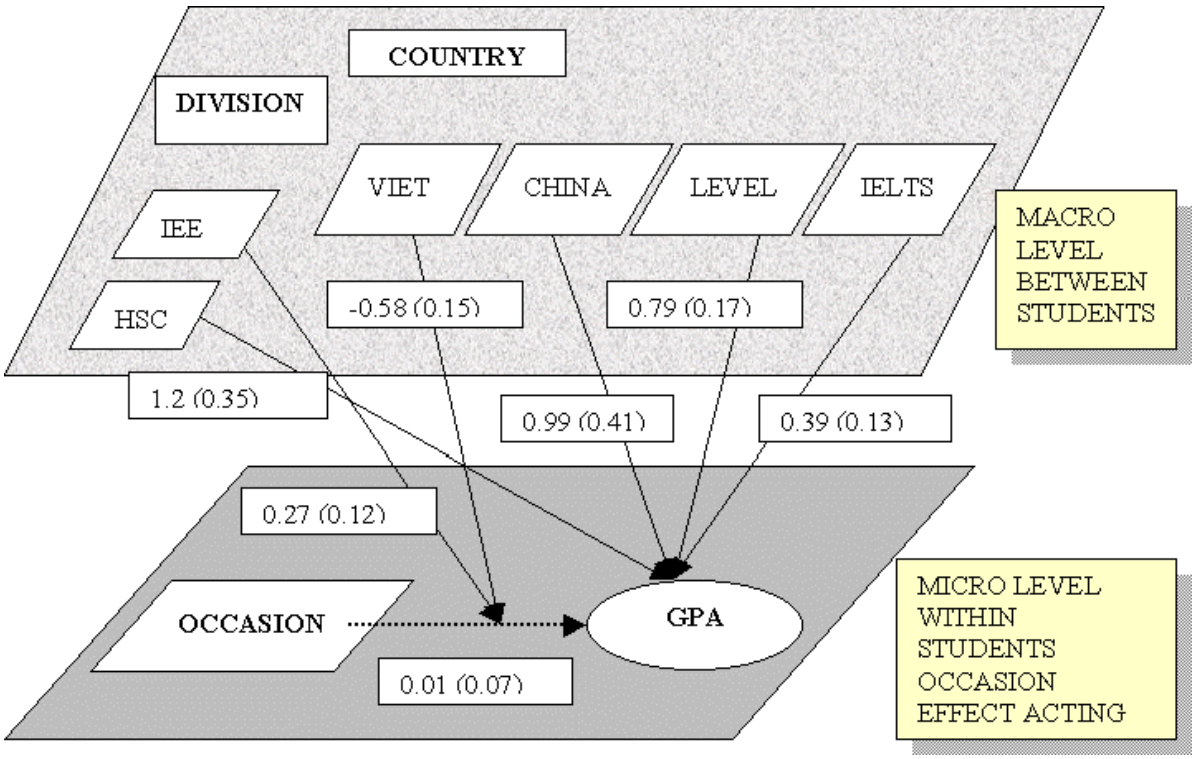
The variable, HSC (7 students who are from the Health Sciences division), has a p-value of 0.001 and a regression coefficient of +1.20 (with a t-ratio of 3.48), indicating a medium level positive relationship between students from the Health Sciences division (with IELTS scores) and GPA. This indicates that on average students from Health Sciences have a higher GPA (by +1.2) than students from other non-Health Science disciplines after adjustments are made for other variables. In practice this shows that Health Science students are likely to be relatively more successful than non-Health Science students with the same IELTS score, all other variables being equal.

Figure 2 provides a diagrammatic representation of the Level 1 and Level 2 effects for IELTS students.

ANALYSIS OF THE SECOND QUESTION

This analysis was undertaken by investigating various methods of selection by raising English proficiency entrance scores and comparing the percentage loss of students with the resulting GPA increases for each method of selection. It should be noted that the IELTS student group were analysed separately for postgraduate and undergraduate students for each equivalent method of selection. These results were analysed separately for postgraduate and undergraduate students because of the findings for Question 1, which suggested that the variable LEVEL was significantly related to mean GPA. Undergraduate and postgraduate

students were shown to have different English proficiency characteristics and therefore were examined separately.



(*Metric regression coefficients are recorded for each effect shown, with standard errors in parentheses. The effect associated with OCCASION is shown as a dotted line because the effect is not statistically significant)

Figure 2*. Summary of Level 1 and Level 2 effects for IELTS students

Undergraduate IETLS students

Table 4 compares the trade-off results of five methods of selection by raising the cut-off score for 46 undergraduate students (all students with a recorded overall IELTS score and a recorded mean GPA) from the present level of IETLS 6.0. Five methods of selection by raising the score are used: Method 1: Overall IELTS score at least 6.5, Method 2: Each module score at least 6.0 regardless of overall score, Method 3: Writing and Reading module score at least 6.0 and overall score at least 6.5**,** Method 4: Writing and Reading score at least 6.0 regardless of total score and Method 5: Writing and Reading score at least 6.0 and total score at least 6.0**.** These methods of selection are chosen because they are ways of minimally raising either the total IELTS score or the module or sub-test scores or a combination of the two. Given the results in Question 1, which show a relatively weak significant relationship between IELTS scores and GPA, it is not considered reasonable that raising the IELTS scores by greater amounts should be considered.

Table 4 clearly shows the trade-off between the reduction in international student numbers with IELTS scores and increase in GPA for the five proposed ways of changing IELTS entry cut-off scores for undergraduate students. Raising the overall IELTS score from 6.0 to 6.5 (Method 1) involves a loss of 50 per cent of the student population with a consequent increase of 0.89 per cent in GPA. Raising each module score to at least 6.0 regardless of the overall score (Method 2) reduces the international student population by 41 per cent and results in a 3.1 per cent improvement in GPA. Requiring a score of at least 6.0 on the

Writing and Reading modules and an overall score of at least 6.5 (Method 3) involves a loss of 57 per cent of international students with an IELTS score and results in a grade increase of 2.88 per cent. Method 4 involves a requirement for two module scores (Reading and Writing) to be at least 6.0 regardless of total scores, which results in a GPA increase of 2.88 per cent and a loss of 39 per cent of students. The final method, which requires a total score of at least 6.0 and a Reading and Writing score of at least 6.0, results in a GPA increase of 0.89 per cent with a loss of 43 per cent of students.

Table 4. Comparative results from raising minimum cut-off scores for undergraduate IELTS students

Five proposed entry Methods	GPA of original students (n = 46)	GPA of remaining students	Change in GPA	% improvement in GPA (base = 4.51)	Original numbers of students	Loss of international students not admitted	% reduction in international student population
1: Overall IELTS score at least 6.5	4.51	4.55	+.04	0.89	46	23	50
2: Each module score at least 6.0 regardless of overall score	4.51	4.65	+0.14	3.10	46	19	41
3: Writing and Reading module score at least 6.0 and overall score at least 6.5	4.51	4.64	+0.13	2.88	46	26	57
4: Writing and Reading score at least 6.0 regardless of total score	4.51	4.64	+0.13	2.88	46	18	39
5: Writing and Reading score at least 6.0 and total score at least 6.0	4.51	4.54	+0.04	0.89	46	20	43

Given the above information extracted from Table 4 and assuming these figures are deemed sufficient to warrant an increase in IELTS scores, Method 5 is recommended as the best way of raising the IELTS cut-off score for undergraduate students. Method 5 provides the security of maintaining the existing cut-off score at 6.0 and raises the minimum standards by using a relatively simple requirement of a minimum score of 6.0 for two of the key skills required at university (Reading and Writing), thus making it relatively simple to administer. It also results in a loss of international students in the middle range (43 per cent) in comparison to the other alternative methods.

The other methods are eliminated for the following reasons:

- Methods 1 and 3 are not selected because they result in an unacceptably high loss of international students (50 per cent or above).
- Despite its relatively low loss of students (41 per cent) and highest GPA increase (3.10 per cent), Method 2 is quite complex to administer and involves the possibility that a lower overall score may result.
- Despite involving the lowest loss of students (39 per cent), Method 4 involves a possibility that this would result in a lower minimum overall cut-off score (below 6.0), as there is no minimum overall requirement.

As there is not much difference between Methods 2, 4 and 5 particularly in terms of student loss percentages, and given that Method 5 results in a lower GPA increase than Methods 2 and 4, it may be that either Method 2 or 4 is considered preferable despite the above

recommendation. Method 2 or 4 may be preferred especially if there is no concern for ensuring a minimum overall score.

Postgraduate IETLS students

Table 5 compares the results of six methods of raising the cut-off score for 51 postgraduate students (all students with a recorded overall IELTS score and a recorded GPA) from the present level of IETLS 6.0. Six methods of selection are used: Method 1: Overall IELTS at least 7.0, Method 2: Each module score at least 7.0 regardless of overall score, Method 3: Writing and Reading module at least 7.0 and overall score at least 7.0, Method 4: Overall IELTS at least 6.5, Method 5: Each module score at least 6.0 regardless of the overall score and Method 6: Reading and Writing module at least 6.0 and overall score at least 6.5. These methods of selection are used as they represent ways of increasing the total and module IELTS scores for postgraduate students to a higher level than for undergraduate students. Given the findings in Question 1, which suggest that postgraduate students with an IELTS score have a greater chance of success than undergraduate students with the same IELTS score, it seems sensible that these scores are able to be raised more than for undergraduate students.

Table 5. Comparative results from changing minimum cut-off scores for IELTS postgraduate students

Six proposed entry Methods	GPA of original students (n = 46)	GPA of remaining students	Change in GPA	% improvement in GPA (base = 4.51)	Original numbers of students	Loss of international students not admitted	% reduction in international student population
1: Overall IELTS at least 7.0	5.19	5.65	+0.46	8.87	51	40	78
2: Each module score at least 7.0 regardless of overall score	5.19	5.50	+0.31	5.97	51	32	63
3: Writing and Reading module at least 7 and overall score at least 7.0	5.19	5.80	+0.61	11.8	51	42	82
4: Overall IELTS at least 6.5	5.19	5.38	+0.19	3.66	51	34	67
5: Each module score at least 6.0 regardless of the overall score	5.19	5.32	+0.13	2.50	51	24	47
6: Writing and Reading module at least 6.0 and overall score at least 6.5	5.19	5.41	+0.22	4.24	51	35	69

Table 5 clearly shows the trade-off between the reduction in international student numbers and increase in GPA for six proposed ways of changing IELTS entry cut-off scores for postgraduate students. Raising the overall IELTS score from 6.0 to 7.0 (Method 1) results in a loss of 78 per cent of the international student population with a consequent increase of 8.87 per cent in GPA. In comparison, raising the overall IELTS score to 6.5 (Method 4) results in a 67 per cent student loss and a 3.66 per cent GPA rise. Raising each module score to at least 7.0 regardless of the overall score (Method 2) reduces the student population by 63 per cent and results in a 5.97 per cent improvement in GPA. In contrast, raising each module score to at least 6.0 regardless of the overall score (Method 5) results in a 47 per cent student loss with a 2.5 per cent grade increase. Requiring a score of at least 7.0 on the Writing and Reading module and an overall score of at least 7.0 (Method 3) results in a loss of 82 per cent of international students with an IELTS score and a grade increase of 11.8 per

cent. Finally, Method 6 produces a loss of 68 per cent of international students and a gain of 4.24 per cent in GPA.

Given these comparisons and the tradeoffs results provided in Table 5 and assuming a willingness to raise IELTS scores for postgraduate students based upon these figures, the recommended method is Method 6. Method 6 raises the overall IELTS requirements, provides the security of a minimum overall score and includes an administratively simple way of raising the minimum score for the two key skills required at university. The other methods were not chosen for the following reasons.

- Methods 1 and 3 both involve an unacceptably high student loss of approximately 80 per cent despite giving the highest GPA increases of all six methods.
- Method 2 is complex to administer and results in a relatively high student loss (over 60 per cent)
- Method 4 is less desirable than Method 6 because it produces a smaller grade increase with a similar student loss percentage and does not guarantee that postgraduate students are proficient in both Reading and Writing
- Method 5 produces the smallest student loss (47 per cent) but involves the lower GPA rise of all methods and is quite complex to administer

Eliminating these methods, Method 6 remains as the preferred method although the student loss of 68 per cent may be considered a difficulty with this method. Indeed none of these methods produces an ideal solution. Indeed Methods 2, 4 and 6 produce relatively similar trade-off results.

Ultimately for both undergraduate and postgraduate student groups, political decisions and processes will decide which methods, if any, the university chooses to adopt. It may even be decided that the small grade increases are not worthy of the relatively large loss of students for both the postgraduate and undergraduate student body and consequently no changes will be made to English language entry requirements on the basis of these figures. It may be considered preferable to consider other ways of strengthening English language proficiencies of international students such as better supporting them once they have gained entry into university. Diverting resources towards employing extra staff skilled in assisting students to improve their English may be a preferred option. However, a mitigating factor is that some of the large losses are due to the fact that students with IELTS scores of less than the minimum score of 6.0 are being admitted to the university. The inclusion of these students (5 undergraduates and 18 postgraduates) as losses in the calculations inflates the loss percentages considerably, especially for postgraduates. Therefore, a first step may be to tighten up current English proficiency entrance procedures to prevent this from happening.

CONCLUSION

Analyses of the results of Question 1 suggest that that there is a positive relationship of IELTS, LEVEL, CHINA and HSC with mean GPA. A strong conclusion is that that there is a significant and positive, but weak, relationship between English language proficiency, as measured by the IELTS scores, of international students and their performance, as measured by their GPA.

These findings complement an investigation of a second question that involves scrutinising various methods of raising the common cut-off scores of IELTS 6.0 for both undergraduate and postgraduate students. Of five methods investigated and on the basis of a trade-off

between the loss of student numbers and GPA gain, for undergraduate students, it is recommended that the best method of raising the IELTS requirements is that the overall IELTS score be kept at 6.0 but a new requirement be introduced to stipulate that students have a Reading and Writing module score of a minimum of 6.0. This result involves a trade-off loss of just over 40 per cent of international students for a GPA gain of 0.9 per cent for undergraduate students. Of the six methods investigated for postgraduate students, it is recommended that the best method of raising the IELTS requirements is that the overall score be raised to 6.5 with an additional requirement of 6.0 in Reading and Writing. This result involves a trade-off loss of almost 70 per cent of international students for a GPA gain of just over 4 per cent for postgraduate students. It is acknowledged that these recommendations may result in unacceptably high losses of international students for very modest GPA gains. Comparatively large reductions in the percentage of students admitted are shown to be the result of raising the minimum IELTS entry scores and thus from a political and financial perspective the university may not choose to raise these entry standards. Although, it is acknowledged that a mitigating factor is that some of these losses are partly attributed to the fact that some students have entered the university with scores of less than the minimum IELTS cut-off of 6.0 or TOEFL 550, a better choice may be to raise support levels for those students who gain entry at present test score levels. This option may assist international students to improve their English language proficiency skills.

Alternatively, linking the two research questions may lead to another conclusion. The multilevel IELTS regression analysis shown in response to Question 1, indicates that if the cut-off score of 6.0 were to be raised to 6.5 for all international students the expected change in GPA would be to lift the mean level of GPA by half of the IELTS score coefficient of +0.39 or 0.20, with all other significant factors remaining unchanged. Moreover, since the analysis of Question 1 also shows that the postgraduate GPA score is already expected to exceed the undergraduate score by 0.80, it may be unnecessary to develop a separate method for changing postgraduate IELTS entry scores. However, given that the analysis includes students with IELTS entry scores below 6.0, it may also be concluded that raising the IELTS entry score for all international students from 6.0 to 6.5 may not be required. Instead, merely not allowing students entry with IELTS scores below 6.0 may be the solution.

This research has merely canvassed some options using statistical methodologies. The final decision, to be made by university management, also involves political and financial considerations.

REFERENCES

Australian Vice-Chancellors' Committee (AVCC). 2001. *Fact Sheet 2: The sources of University Income*, Australian Vice-Chancellors' Committee. [Online].] Available: http://www.avcc.edu.au/ [2001, 3rd December].

Bellingham, L. 1993. The relationship of Language Proficiency to Academic Success for International Students. *New Zealand Journal of Educational Studies*, 30 (2), 229-232.

Burns, R. 1991. Study and Stress among First Year Overseas Students in an Australian University. *Higher Education Research and Development*, 10 (1), 61-77.

Byrk, A. S. and Raudenbush, S.W. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*, California: Sage Publications.

Charge, N and Taylor, L.B. 1997. Recent Developments in IELTS. *ELT Journal*, 51, (4), 374-380.

Ciccarelli, A. 2001. IELTS issues for UniSA. Unpublished paper for Senior Management Group, University of South Australia, Adelaide.

Cotton, F. and Conrow, F. 1998. An Investigation of the Predictive Validity of IELTS amongst a Group of International Students studying at the University of Tasmania. *English Language Testing System Research Reports*, 1, 72-115.

Criper, C. and Davies, A.1988. IELTS Validation Project Report, The British Council, University of Cambridge Local Examinations Syndicate, London.

Dooey, P.1999. An investigation into the predictive validity of the IELTS Test as an indicator of future academic success. In Martin, Stanley and Davison (eds) *Teaching in the Disciplines/Learning in Context*, 114-118, Proceedings of the 8th Annual Teaching Learning Forum, University of Western Australia, February 1999.

Elder, C. 1993. Language proficiency as a predictor of performance in teacher education. *Melbourne Papers in Language Testing*, 2 (1), 68-87.

Ferguson, G. and White, E. 1993. A small-scale study of predictive validity. *Melbourne Papers in Language Testing*, University of Edinburgh, 15-63.

Fiocco, M. 1987. English Proficiency Levels of Students From a Non English Speaking Background: A Study of IELTS as an Indicator of Tertiary Success. *Research Report*, Centre for International English, Curtin University of Technology.

Gibson, C. and Rusek, W. 1992. The validity of an overall band score of 6.0 on the IELTS test as a predictor of adequate English language level appropriate for successful academic study. Unpublished Masters of Arts (Applied Linguistics) thesis, Macquarie University, New South Wales.

Gue, L.R. and Holdaway, E.A. 1973. English Proficiency Tests As Predictors Of Success In Graduate Studies In Education. *Language Learning*, 23 (1), 89-103.

Graham, J.G. 1987. English Language Proficiency and the Prediction of Academic Success. *TESOL Quarterly*, 21 (3), 505-521.

Hill, K., Storch, N. and Lynch, B. 1999. A Comparison of IELTS and TOEFL as Predictors of Academic Success. *English Language Testing System Research Reports*, 2, 52-63.

Ho, D.Y.F. and Spinks, J.A. 1985. Multivariate Prediction of Academic Performance by Hong Kong University Students. *Contemporary Educational Psychology*, 10, 249-259.

Hox, J.J. 1994. *Applied Multilevel Analysis*. TT-Publikaties, Amsterdam.

IELTS. 2001. *International English Language Testing System Handbook*, joint publication of UCLES, the British Council and IDP Education Australia: IELTS Australia**.**

Kerstjens, M and Nery, C. 2000. Predictive Validity in the IELTS Test: A Study of the Relationship Between IELTS Scores and Students' Subsequent Academic Performance. *English Language Testing System Research Reports*, 3, 85-108.

Light, R.L., Xu, M. and Mossop, J. 1987. English proficiency and academic performance of international students. *TESOL Quarterly*, 21 (2), 251-261.

Nelson, The Hon. Dr B. 2002. *Setting Firm Foundations: Financing Australian Higher Education*. Department of Education, Science and Training, Canberra.

Rusek, W. 1992. IELTS: Does It Predict Success At University? Proceedings of the 5th Annual Education Conference, Adelaide, 1-3rd October.

Tatsuoka, M.M. 1997. Regression Analysis of Quantified Data. In J.P. Keeves (ed), *1997 Educational Research, Methodology, and Measurement: An International Handbook*, 2nd edition, Cambridge University Press, Cambridge, 648-657.

Traynor, R. 1985. The TOEFL: An appraisal. *ELT Journal*, 39, 43-47.