

International Education Journal



Volume 7 Number 5
October 2006

Published by

Shannon Research Press
Adelaide, South Australia
ISSN 1443-1475
<http://iej.com.au>

International
Education
Journal



Volume 7 Number 5
October 2006

Published by

Shannon Research Press
Adelaide, South Australia
ISSN 1443-1475
<http://iej.com.au>



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, formal and informal, 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, free of charge.

Peer Review - All articles appearing in this journal have been subjected to a blind peer-review process by at least one expert in the field.

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

Publication Frequency - It is intended that the journal be published four 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.

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

Published by:

Shannon Research Press,
39 Dorrien Avenue,
Woodcroft, SA 5162,
Australia

Copyright © 2006

Shannon Research Press

ISSN 1443-1475

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

Designed by Katherine Dix

Printed in Adelaide

October 2006

CONTENTS

To be fat or thin? Social representations of the body among adolescent female students in Brazil Stenzel, L.M., Saha, L.J. and Guareschi, P.	611
Using regression analysis to establish the relationship between home environment and reading achievement: A case of Zimbabwe Kanyongo, G.Y., Certo, J. and Launcelot, B.I.	632
What makes a difference between two schools? Teacher job satisfaction and educational outcomes Lee, M.	642
Why do they not talk? Towards an understanding of students' cross-cultural encounters from an individualism/collectivism perspective Tan, J.K.L. and Goh, J.W.P.	651
Comparing university academic performances of HSC students at the three art-based faculties Ismail, N.A. and Othman, A.	668
Private education as a policy tool in Turkey Cinoglu, M.	676
Does increasing communication through visual learning environments enhance student perceptions of lecturers? Frumkin, L.	688
Individual and flexible: Working conditions in the practice of Swedish distance-based teacher education Lindberg, J.O. and Olofsson, A.D.	699
State educational policy and curriculum: The case of Palestinian Arabs in Israel Abu-Saad, I.	709
Insisting on equity: A redistribution approach to education Nordstrum, L.E.	721
Retention and academic achievement research revisited from a United States perspective Lorence, J.	731
Successful pedagogies for an Australian multicultural classroom Winch-Dummett, C.	778

Editorial Board

Flinders University
Bedford Park, South Australia
GPO Box 2100, Adelaide 5001
Fax: 08-8201 3184

Editors in Chief

Prof John Keeves
john.keeves@flinders.edu.au
Prof Jonathan Anderson
jonathan.anderson@flinders.edu.au

Online Editor

Ms Katherine L. Dix
katherine.dix@flinders.edu.au

Executive Assistant

Dr Helen Askill-Williams
helen.askell-williams@flinders.edu.au

Assistant Editor

Dr Tony Gibbons
tony.gibbons@flinders.edu.au

Book Review Editor

Dr Abul Farooque
abul.farooque@flinders.edu.au

ERC Assistant Editor

Ms Mirella Wyra
mirella.wyra@flinders.edu.au

ESL Editor

Dr Bobbie Matthews
Flinders University Institute of
International Education
bmatthews@internode.on.net

Editorial Advisory Board

Prof Bob Catley
Department of Political Studies
University of Otago
Prof Kevin Marjoribanks
Graduate School of Education
University of Adelaide
Dr Larry Saha
Department of Sociology
Australian National University
Dr Wendy Duncan
Education Specialist
Asian Development Bank, Manila

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 by email to:
onlineeditor@iej.cjb.net

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.

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.

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 endnotes, 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

To be fat or thin? Social representations of the body among adolescent female students in Brazil¹

Lucia Marques Stenzel²

Foundation Federal Faculty of Medical Sciences of Porto Alegre Brazil lucia@stenzel.com.br

Lawrence J. Saha

The Australian National University lawrence.saha@anu.edu.au

Pedrinho Guareschi

Pontifical Catholic University of the State of Rio Grande do Sul Brazil guaresch@pucrs.br

The aims of this paper are (a) to investigate how adolescents perceive and represent the body form with respect to being fat or thin, and (b) to describe the process of how they constructed the social representations for these latter two body conditions. The data were collected by means of individual and focus group interviews with adolescent female students in Brazil who were from 11 to 21 years of age. When the adolescents were questioned about their bodies, they talked about 'being fat' or 'being thin', even though they were not asked about weight issues. Following their own logic, they did not portray 'feeling fat' and 'feeling thin' as related to their 'real' body condition or weight. Furthermore, in the adolescents' discourse, the concept of 'normal weight' was virtually non-existent and was characterised as 'nothing' or 'more or less'. By the end of the interviews, their depictions of these conditions of body weight included links from the body to their social relationships in the form of perceived group exclusion or inclusion. In our discussion we describe the adolescents' collective discourses on being fat or thin as integrated social representations, which incorporate both the physical and interpersonal dimensions of their experiences. We conclude by examining the practical implications of our findings for female adolescent behaviour, especially with respect to obsessive dieting and possible eating disorders. Finally, we explore the possibility of educational programs to counter the media and other influences which give rise to the negative aspects of social representations of the body by adolescents.

Social representations, female, adolescents, body image, body weight, health education, Brazil

INTRODUCTION

Today, being overweight is generally considered to be unhealthy and not attractive. However, this was not always the case, as being overweight was once considered to be a characteristic of a healthy or beautiful body. Even across cultures, the valued forms of the body, obese or slender, do differ considerably in terms of health or beauty (Garine and Pollock, 1995; Nasser, 1997; Stearns, 1997).

¹ An earlier version of this paper was presented at the VII International Conference on Social Representations, Guadalajara, Mexico, 10-14 September, 2004. We wish to acknowledge the valuable comments made on an earlier version of this paper by Gary Dworkin (University of Houston) and John Keeves (Flinders University).

² Lucia Marques Stenzel was a postgraduate student at the Pontifical Catholic University of the State of Rio Grande do Sul (PUCRS) when the data were collected. She was a Visiting Fellow from September 2003 to September 2004 in the School of Social Sciences, The Australian National University, when this paper was written. She was a recipient of a CAPES scholarship from the Brazilian government during this period.

In modern Western societies, there has emerged a culture of thinness whereby being thin has become a goal to be pursued for both aesthetic and health reasons. On the other hand, being fat or obese has become a condition to be avoided. Social scientists, anthropologists and historians point out that our perception of the body is a social and cultural construct that depends on the historical and the social context (Sobel and Maurer, 1999). These perceptions of the body are transferred to youth through the media, and have a marked influence on their lives in schools and universities through the activities of the peer cultures that exist in educational institutions.

Body weight is an issue of particular relevance for adolescents in many societies. Not only is it related to their body image and self-identity, but they are particularly susceptible to the negative health consequences of bodyweight extremes, such as obesity, or anorexia and bulimia for those preoccupied with being thin. Preoccupations about weight, diet, and body form are a part of contemporary adolescent culture, and do not affect only the adolescent group that is overweight. The preoccupations about weight affect especially young female adolescents. Dieting is thus one response to the pressure to conform to the ideal image of women to be thin (Chernin, 1981; Fallon, Katzman and Wooley, 1994; Germov and Williams, 1999; Orbach, 1978; Paquette and Raine, 2004; Székely, 1988; Wolf, 1990). Some writers argue that in general women tend to be dissatisfied with their bodies (Stinson, 2001) even though their bodies are objectively “normal”. Some young women actually dislike their bodies, a phenomenon called “body hatred” (Frost 2001). Body weight is a large part of this dissatisfaction. Frost calls it the “the thin imperative” (Frost, 2001, p.196) In its extreme form, this phenomenon, sometimes clinically labelled as Body Image Disturbance, can lead to disordered behaviours such as crash dieting, fasting, bingeing, purging and more dangerous eating disorders (Posavac and Posavac, 2002).

Preoccupations with weight and appearance have been extensively examined in psychiatric and medical research, especially in the eating disorders field (Bruch, 1962; Gordon, 1990). The connection between the concept of body image and eating disorder problems made an important impact on the studies of how individuals subjectively perceive their bodies, although it has also reinforced a limited notion of body image (Cash, 2004). Body weight concerns are not limited to clinical populations. ‘Feeling fat’, for example, is a common complaint voiced by females of all ages in Western societies; “... a theme of dissatisfaction is evident in the research on body image among women without eating problems” (Haworth-Hoeppner, 1999, p.91). These issues have been documented, mainly in highly developed Western societies. This article draws on evidence from a society that is a part of Western culture, and indicates that also in a Latin American culture youth are greatly concerned about their body image (Nasser, 1997).

Female adolescents in Brazil are an example of the above condition, as they are subjected to a cultural environment which puts high priority on body appearance. This cultural pressure has been documented in a study by Nunes and her colleagues in Porto Alegre (Nunes, Olinto, Barros and Camey, 2000; Nunes, Barros, Olinto, Camey, and Mari, 2003). They found the following in their research:

...forty-six per cent of the women had an ideal weight lower than their actual weight and 37.8% considered themselves as fat. Among women with normal BMI, 25.2% who were classified as normal, presented risks of abnormal eating behaviour and 5.7% had eating disorder symptoms. Among women that considered themselves fat, 47.2% presented risk behaviour and 19.2% had eating disorder symptoms” (Nunes et al., 2000, p.2, translation ours).

We want to arrive at a deeper understanding of the quantitative results reported by Nunes (Nunes, et al. 2000; Nunes, et al. 2003). In particular, we want to know the cause of the feeling that adolescent girls in Brazil have of being overweight without really being overweight. As Cash (2004) has pointed out, “... individuals’ own subjective experiences of their appearance were

often even more psychosocially powerful than the objective or social “reality” of their appearance” (p. 1).

In order to explain this phenomenon, we will focus our attention on the underlying social representations of the female body, that is, the socially constructed meanings that young Brazilian women have about their bodies, and consequently about being fat or thin. Thus, we are concerned with how and why young adolescent women, at least in Brazil, disregard the objective weight condition of their bodies and regard them in a way that is consistent with these cultural pressures.

In focusing our inquiry on adolescent women, we do not intend to imply that the constructions of the body occur in a vacuum. We are aware that the representations which they construct must be seen in a context, namely that they are students in schools, they are members of peer groups, and that they are exposed to influences from the wider society, especially the media. While this fact does not deter us from focusing on the construction of the social representations as a process, the interpretation of these representations must take the context into account. We discuss these implications of the constructed representations at the end of this paper.

THE BODY AS SOCIAL REPRESENTATION

We will use social representation theory to investigate the meanings that adolescents have about their bodies, body weight and body image. The current concept of social representation has its origin in the writings of Moscovici who originally developed the theory in his analysis of the diffusion of psychoanalytic concepts in French society (Moscovici, 1976). Social representations are an essential part of communication, and as such also can be defined as the “manner in which values, ideas, and practices are structured in and by ordinary communication, allowing people to both communicate and to order their world” (Semin, 1995, p.601). They emerge from conversations in which people attempt to make sense of their social lives (McKinlay and Potter, 1987), or as Moscovici would say, in these conversations people are “creating reality” (Moscovici, 1987; p.517). These social representations serve as “reference points” which make communication possible with another person. In other words, social representations consist of shared knowledge. As Duveen puts it, “representations are the result of communication, but without representations there would be no communication” (Duveen, 2000, p.4).

Social representation theory directs attention to features of daily discourse about a particular idea, value or practice. Because social representations are socially and culturally shared, the search for their ultimate origins often leads to the consideration of agents such as the media, parents, peer groups, and schools, depending on the target group being studied. However, in the course of communication the objects or events from these agents become new ‘common-sense knowledge’ which is shared and diffused at a new level.

In our case, this discourse is about the characteristics of the body, and how it is related to weight. Therefore the object of our study is the way that groups of Brazilian female adolescents manifest shared understandings about the body and body weight in their daily discourse. We will focus on the content of the discourse, and what happens to the content in the process of discourse.

Thus when we investigate the social representations of the body for Brazilian adolescent girls, we want to know both the *content* and the *process* by which the representations emerged. In other words, we want to know, how the content was arrived at, or the “familiarisation of the unfamiliar” (Semin, p.103; also McKinlay and Potter, 1987; and Moscovici, 1987). In this way, we will not only acquire the knowledge of the attributes or the characteristics of the body which make up the discourse of these adolescent females, but also how these representations of the body come to be understood by them.

According to social representation theory, it is this shared knowledge about the body and its attributes which makes communication possible among a particular group of persons. Also, our knowledge of how the body is understood by these female adolescents helps us understand the pressures that they feel, and the behaviours which are related to them.

METHOD OF THE STUDY

The choice of method in social representation research is not dictated by the theory itself. Actual research studies using social representation theory have included both qualitative and quantitative approaches, and both experimental and survey designs. As Bauer and Gaskell comment, "...almost every method known to the social sciences has been used in the study of social representations" (1999, p.163). Our data were obtained through the use of focus group interviews.

The Data

Focus group interviews were conducted in 2000 with groups of young women between 11 and 21 years of age, who were students from Grade 5 to university, in Porto Alegre, Brazil. The selection of age and grade range was made to take into account their diverse experiences and vulnerabilities with respect to eating disorders. In total, four focus groups, with a total of 25 participants, were formed with adolescent young women of roughly the same age and grade. This is consistent with Morgan (1997) and Bloor, Frankland, Thomas and Robson (2001) who argue that three to five focus groups are usually sufficient to attain "saturation" in the data, that is, that no more "meaningful new insights" would likely occur by increasing the number further (Morgan, 1997, p.43).

The focus group interviews lasted from between 45 minutes to one hour and 15 minutes. They were conducted by the first author of this paper. In addition, two in-depth interviews with individuals were conducted. The first individual interview served as a pilot in order to develop the structure for the focus group interviews. The second interview was conducted after the focus group interviews, and served as a validity check on the focus group interviews and made possible the exploration in greater depth of some of the topics raised in the focus groups. In the analysis, all interview data, both from the focus groups and the individuals, are considered of equal value and are used to illustrate the construction of the social representations. In our analysis, the discourses of the focus group interviews and the individual interviews were cross-validated, and this is the basis for regarding both sets of data as of equal value.

All interviews were tape-recorded, and the tapes were transcribed in their entirety, as usually recommended (Stewart and Shamdasani, 1990; Bloor, et al, 2001) and transformed into informal texts, keeping the order of the themes as they emerged in the discourse. In order to describe and analyse the findings, the underlying goal of the project was kept in the forefront, namely the understandings of the participants about the body. The procedure entailed the identification of speech marks, or quotes that related to the question underlying the project itself, namely what is the social representation of forms of the body for these young women? This question was not used to begin the interviews because one of the goals of the project was to confirm if and when the theme would emerge spontaneously. So the first question of the interviews only introduced the topic of the body, as forms of body weight were considered attributes of the body and were left for the participants to bring up. Only after these ideas were mentioned did the interviewer facilitate the discussion. As a result, the process by which the topic of being 'fat or thin' eventually emerged in the discourse of these adolescents also is a part of the analysis and findings. The fact that some ideas were mentioned at an early stage of all interviews represents an important finding in its own right which will be discussed later. Therefore, through this procedure the deeper meanings and understandings of body weight held by the participants, were uncovered.

All interviews began with the same question, namely, ‘What comes to your mind when you hear the word body?’ There were no structured questions following the initial question.

Analysing the Data

The analysis is built around three fundamental concepts, namely themes, categories and units of analysis (Spink, 1995). The operational definitions of these concepts are given below. In accordance with the primary feature of focus group methodology, namely the interaction between members of the group, we have chosen to include passages which are indicative of this interaction as the social representations of the body emerge in the interview texts (Morgan, 1997; Bloor, et al., 2001).

First, themes are general concepts which relate to the original problem of the study. These themes are made up of a number of categories. A category, as our second analytical concept, is a group of units of analyses that have common meanings. The third concept, units of analysis, consists of speech marks, or short quotes, sentences, or dialogues between the participants, that contained the same meaning and are added together. These units of analysis varied in length. In the focus groups, for example, sometimes more than one participant was talking about the same idea regarding the topic. Therefore the units of analysis could be a series of single statements, or a dialog between the participants about the same idea.

This analytic procedure can be portrayed in the following manner:

1. Themes: categories with a similar theme,
2. Categories: units of analysis with similar meanings, and
3. Units of analysis: speech marks with a common meaning.

We display our findings in an innovative visual and textual manner. The themes and categories emerging from the focus group discussions are represented as concentric rings with appropriate labels. These are accompanied by the relevant ‘speech marks’ or quotes from the focus group discussion which illustrate and give support to the identification of the themes and categories. The concentric rings indicate dynamism, because the categories are dynamic and not static. This means that each ring represents one category and includes its two opposites. The positive and negative poles portray the dialectical nature of the category in the discourse. This visual display technique will become more clear when we present the first figure of our analysis.

FINDINGS

In this section we map out the meanings that emerged in the course of the focus groups, starting with the original theme contained in the question of the interviewer: “What comes to your mind when you hear the word body”. The purpose of this procedure is to present in a visual manner the process by which the social representations were constructed by the focus group participants (Puddifoot, 1997). In this way, we also make explicit the way the adolescents understand and give meaning to the weight conditions about which we spoke at the outset of this paper.

Theme 1: The Body

The first theme, the body, was the one that was proposed initially by the researcher. After the original question was asked, the adolescents began their discussion.

Category 1: My body and the body of others

The first category that emerged from the discussion was the distinction between ‘my body’ and ‘the body of others’. This distinction is reflected in the following ‘speech marks’ which were

taken from the individual interviews and relevant focus group interview. This procedure of data presentation will be followed in all subsequent illustrations taken from the interview material.

Well... (pause) what comes to my mind....(pause). My body. (Interview 1)

Until I turned 13, I did not really care about my body, as I was chubby (pause). Then, after I got my first period (menstruation) and my body started to change, then I began to think about my body ... (Interview 2)

Body? Mmmmm..., let me see... I don't know! Maybe looking at others..., the way they look. I mean looking at a man (everyone in the group starts to laugh). Of course! You guys think that I should look at women's bodies? (Participant 1, Group 3)

Although the adolescents start talking about the bodies of others, the content of their discourse is mainly about their own bodies. However, they talk about their own bodies in association with the bodies of others in order to make distinctions and to define better their own bodies.

Why do these adolescents start talking about 'their bodies' if the stimulus was only the word 'body'? If we go back to the literature about adolescents we find that this first reaction to the stimulus is not new, but even predictable. Adolescence is a time for a rapid and fundamental physical change and it is natural that adolescents are conscious of, and talk about their bodies and their experiences with these changes (Coleman and Hendry, 1999). However, the next category that emerges in association with that of 'my body' and 'the body of others' is that of 'physical appearance'. This category emerged at a very early stage in some groups, like the speech marks exemplified before, and gives an important specific meaning to the previous dialectical category.

Category 2: The physical appearance of the body

The meaning that adolescents give to the original theme (body) is related to the second category that we call 'the physical appearance of the body'. The participants could have talked about a range of other aspects of the body, such as the function of the body, or the health of the body, or even the symbolic or spiritual meanings of the body. But the discourse was very clearly focussed on the appearance of the body.

When you are a teenager..., (pause) I mean..., when you turn 13, you start to think a lot more about the way you look and the way the others look like (Participant 1, Group 2)

I think image is everything! People can say that that they don't really care about appearance, but the physical image is everything! Well..., at least at the first moment of contact. (Participant 3, Group 2).

In the process of the interview, each category that emerges gives more specific meaning to the previous categories. Therefore the category that we have labelled 'appearance' did not occur without itself being further elaborated. As the interviews progressed, the participants became more precise about their understanding of the theme. And therefore at this stage of the discourse, the adolescent girls started to differentiate between the 'problematic' appearance and the 'ideal' appearance of the body.

Categories 3 and 4: Problems of the body and the ideal body, and fat and thin

This leads us to the third category that emerges from our analysis of the discourse, namely 'Problems of the Body' and 'The Ideal Body'. These dialectical categories express the adolescents' expectations of the way, in their own eyes, the body should appear in association with what they do not appreciate in their own bodies. They immediately discuss their notion of the perfect body, and give examples of people that they believe have perfect bodies. The

'Problems of the Body' are related to their own bodies, that is, what they do not appreciate about their bodies, what they dislike in their bodies, the imperfections they see in their bodies, and what they would like to change in their bodies.

I think about the problems of my body. I think I have too much hair on my body, that I'm too hairy ... sometimes I think that my boobs are not big enough, ... sometimes I think that I'm fat ... this kind of stuff (Interview 1)

... Someone could think about a perfect body, or about their own body. But me...I don't know.... (Participant 3). The body changes with time (Participant 1). Keeps changing... (Participant 2); It's a shame that it (the body) starts to look...not that good. Maybe later, if you have problems and you start to dislike (the body) you can always go on a diet, or do another kind of intervention... then it gets better (Participant 1, Group 2)

I have a lot of problems with my body. I hate my body (the group reacts and starts talking at the same time). I'm serious! (the participant tries to continue). This summer I decided that I wanted to look like the "Feiticeira"³. I'm serious girls! I don't feel comfortable with my body. If I could be like her... I don't want "to be her" (Feiticeira), but I want to have her body. I wish I could have a nice belly, pretty legs, and a nice bum...(Participant 1, Group 4)

The progression of the discourse at this stage of the analysis is visually represented in Figure 1.⁴

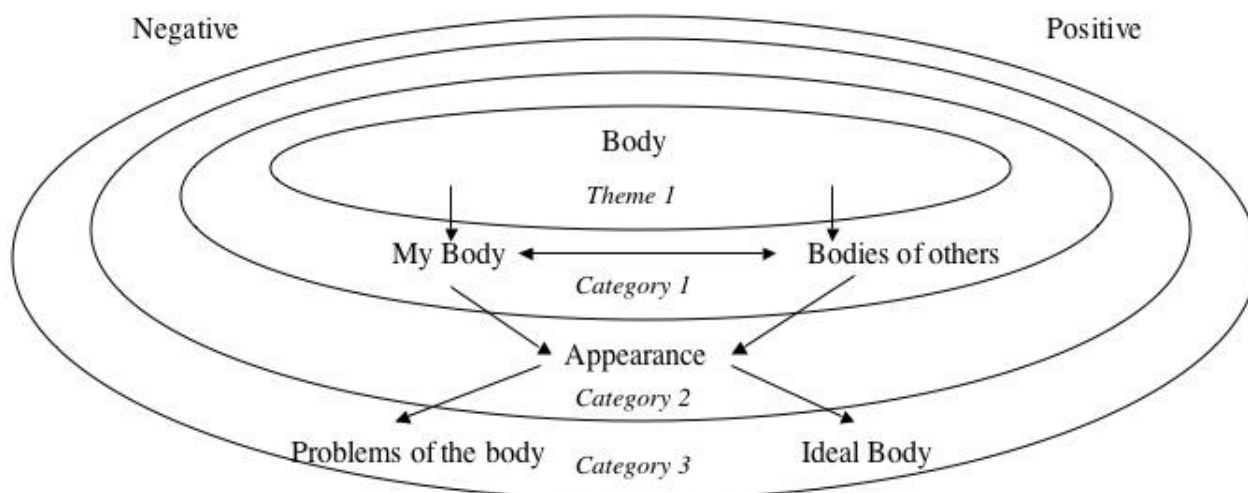


Figure 1. The analysis of the discourse through the first three categories

What we can observe in the course of the analysis thus far, is that, on the one hand, the speech marks that bring together the meaning of the category 'Problems of the Body' are related to the participants' own bodies. On the other hand, the bodies of others are exemplified as the ideal body. The speech marks illustrate the discontent of participants with their own bodies.

... When I take my clothes off and I look at myself in the mirror, I start to check my body. I turn to one side and the other and I usually dislike things, like my belly for example... (Interview 1)

³ Feiticeira is a Brazilian TV icon, who, by 2000, was considered by the media to be the most beautiful woman, and who had the most "perfect body" in Brazil. Many magazines have had her picture on their covers.

⁴ We have included the label "category" in Figure 1 in order to better illustrate the analytical structure of the diagram. In order to simplify subsequent figures, we have omitted this label.

I hate my nose! (the group starts laughing) That's because it is flat. I don't like my body at all... (Participant 2); That's horrible! (Participant 3); I don't like my hair (Participant 8); Me too! I hate my hair! (Participant 3); Well..., my hair... (Participant 2-reflexive moment); I don't like my face (Participant 3); Sometimes I like my hair..., but only sometimes (Participant 4); I don't like my hair either. I don't have enough hair (Participant 8); Yes you do have a pretty hair! (Participant 1). No (I don't)...(Participant 8, Group 1)

... I wish I could have a perfect body. I weigh 63 kilos now, but I want to go down to 58 kilos. ... I'm doing a lot of exercise because I want to go down to 58 kilos. I'm worried... Sometimes I say: "I don't care!" But of course I do. I look at myself in the mirror and I don't like what I see (Interview 2)

At this point, our analysis helps to clarify what the category 'Problems with the Body' means, and we get closer to the focus of our study. As was previously mentioned, the discourse about the body and its categories, namely appearance, problems with the body, and the ideal body, move spontaneously to the concepts that they called 'fat and thin'.

We use these concepts to label the next theme in the discourse, namely 'fat' and 'thin'.

Theme 2: Fat and Thin

When the young girls started to mention the theme 'Fat and Thin', the interviewer began to direct the conversation in order to help them express their understandings about this second theme. The following example is taken from Group 4:

Group 4: ... sometimes I don't even look at myself in the mirror, because if I do I'll think that I'm so ugly (Participant 2) ...And then you don't want to go out of the house (Participant 5).

Interviewer: So that means that you are not happy with your body?.

Group 4: I am! (Participant 2); I am not! (Participant 1). Well... I can't say that I'm not happy with my body, but... I think I need to lose weight (pause). Sometimes I am happy and sometimes I'm not.. Well..., I guess I can say that I'm not really happy with my body (Participant 3). I also have this problem... But since I started to lose weight I started to feel better with my body (Participant 4).

Interviewer: Let's talk a little bit more about this, about losing weight and feeling better about your body. I want to understand better what you mean by this.

Figure 2 represents the process of the discourse and the change of theme as stimulated by the interviewer.

At this point the adolescent girls began to focus on the theme of 'Fat and Thin' by trying to conceptualise what it is to be fat or thin from their perspective. Even though their focus shifts to a different theme, it is important to observe that there is continuity between the previous theme and categories.

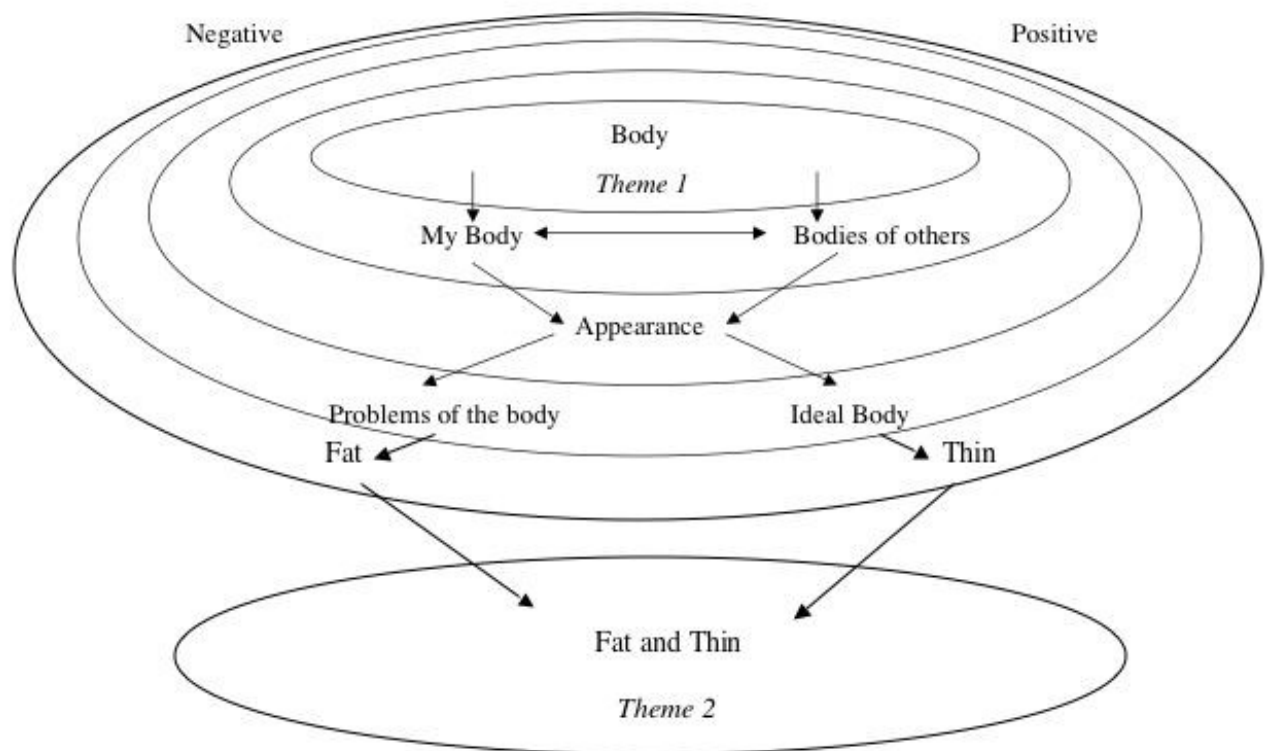


Figure 2. The emergence of the second theme, Fat and Theme

Categories 1 and 2: Being fat and being thin, and feeling fat and feeling thin

When the theme of being fat or being thin emerges, two different categories appear in the discourse at once. When the adolescents start to define what they call 'fat' and 'thin', the definitions are articulated in two different ways, namely 'being fat' and 'feeling fat', and 'being thin' and 'feeling thin'.

'Being' and 'feeling' are two different realities when the focus is on body weight. Feeling fat or thin does not have to do necessarily with being fat or thin. In other words, their perceived body weight does not correspond to their objective body weight. This distinction is an essential finding of this study, not only because it shows that 'feeling' is different from 'being', but because of the symbolic meanings that underlie this separation. The important question is what makes them feel fat without really being fat?

We argue that these adolescents are reflecting a different logic that guides their understanding and consequently their behaviour regarding body weight. This logic, as well as the distinction between 'being' and 'feeling' is represented in Figure 3. The broken arrows represent the indeterminate logic of the adolescents that will become more clear and important later in this analysis.

The complexity of these distinctions and the interrelations of the categories 'Being Fat or Thin' and 'Feeling Fat or Thin' are exemplified in the selection of speech marks found below:

For me it is like this... Some of them are too thin, others are a little bit fat, others are normal, you know? But all of them keep saying that they are fat. The ones that are thin want to keep losing, and losing, and losing weight! (Interview 1)

... I don't like it when a thin person keeps saying: "I'm fat! I have to go on a diet!". They say this because they want to hear compliments about their body..., like: "Oh gosh! You are so skinny!", you know? (Participant 2). Well..., I sometimes do that...

But that is not my intention... My intention is... (she pauses and reflects)... because, sometimes I look in the mirror and I really feel fat...(Participant 1) My mom lost a lot of weight, I think more than 20 kilos or maybe even more. But she keeps saying that she is fat. She does not see herself as a thin person. (Participant 3; Group 2)

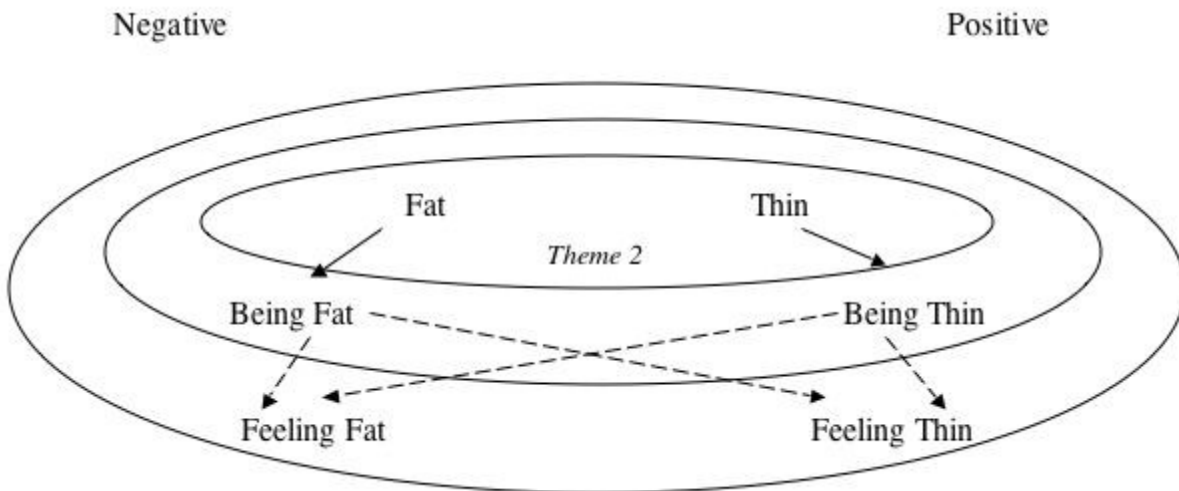


Figure 3. The logic of adolescents about Fat and Thin

Category 3: Normal

In the above discourse, a new category appears, namely that of 'normal'. However, what is interesting is that this category simply disappears as the discourse flows on. But 'Being and feeling fat' continues to receive meaning as does 'being and feeling thin'. 'Being and feeling normal' takes on a neutral meaning. In the words of one of the participants, it is like "feeling more or less". Clearly the category 'Normal' has lost its significance for these adolescents. If 'normal' means feeling 'more or less', then why would an adolescent want to be only 'normal'? The discourse regarding 'being normal' is given in Figure 4.

In contrast to the concept of 'normal' as having a neutral meaning, the categories 'fat and thin' are understood as being at opposite ends of a continuum. The concepts that are used to give meaning to the category 'being fat', for example, are clearly negative. The adolescents use stereotyped negative expressions when describing overweight persons, and they regard them as totally responsible for their weight condition. These views are reflected in the following speech marks:

You are obese if you do not take care of yourself... It's not healthy. (Participant , Group 4)

The person is fat if she or he wants to be fat! (Participant 1, Group 3)

She does not want to take care of herself... (Referring to an overweight cousin)...I always try to advise her...But my cousin is fat because she wants to be. (Interview 2)

...Once my brother saw this woman in the bus and she could not pass through the turnstile ...She was stuck! (The group starts laughing), (Participant 2). I also watched that happen once!! And you know what? Everyone had to get out of the bus through the back door (The group laughs even louder) (Participant 3). The woman was trapped? (Participant 1). (The group keeps laughing aloud). I think they had to cut part of the turnstile (Participant 3). Oh Gosh! What an embarrassing situation! (Participant 1). What a horror! (Participant 8). Thank God she was not your relative! (Participant 1, Group 1)

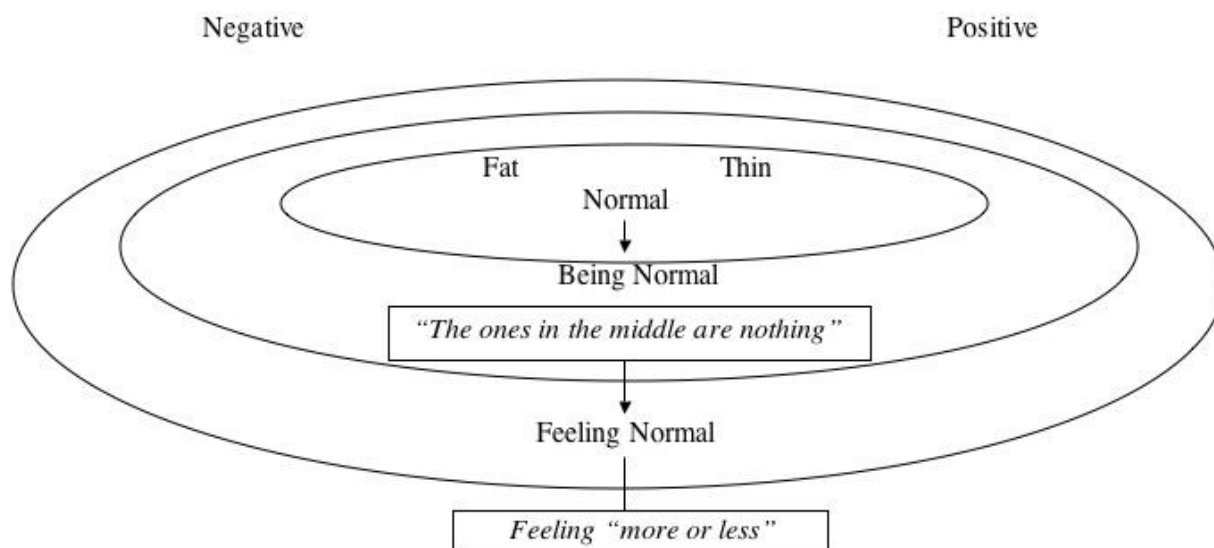


Figure 4. The category, Normal, with the corresponding units of analysis (quotes)

As we see in the speech marks above, these adolescents are inclined to describe situations related to the category ‘being fat’ in a way that is unconnected with their own body weight experience. In this part of the discourse we observe that the adolescent girls gradually become conscious about the way their own experiences and feelings are a part of the phenomenon (being fat or thin) about which they are speaking.

...I don’t know... I put myself in the place of this person who was stuck in the turnstile... Terrible...(Participant 2). They suffer...they are always the objects of jokes ...(Participant 4). Jokes (reflecting)... (Participant 2). Sometimes people even say that they smell bad... (Participant 4). People think they are funny, people think...(Participant 5). Exactly! (interruption) (Participant 4). Because fat people have to have something good, because they are fat... At least they have to be funny to be accepted. (Participant 5, Group 4)

No one likes to be fat...everyone wants to be thin. Ok, ok... Me too... I also think that being thin is better than being fat... Fat people are treated different... They are isolated..., they are left aside... The thin ones are always welcome! (Interview 1)

As we mentioned in our discussion for Figure 3, the feeling of these adolescents about body weight does not correspond to objective body weight. It is a different logic that guides their understanding and consequently their behaviour regarding body weight. The following speech marks illustrate the category ‘Feeling Fat and Thin’, where the adolescents also talk about their behaviour which is based on this logic.

Something that I can observe is that, people of our age are always changing their weight. I am like that. I gain and lose, gain and lose, gain and lose, all the time. Then I feel really bad, because...you gain..., then you lose and everything is perfect! You feel great and happy...then you stop to take care of yourself and everything goes back... and you feel bad again. It’s always like that, this roller coaster: gain and lose, gain and lose, gain and lose... This makes people feel bad...I think... I feel bad... (Participant 3, Group 4)

...I do not think about ‘stopping eating’ because I do a lot of exercise..., I cannot accept the idea of getting fat and that is the reason why I am always on a diet. I cannot imagine myself ‘not dieting’. Weight always worries me... (Participant 1, Group 2)

...When I was about 12 years old I knew the calories of everything. Me and my friend, who was actually fatter than me,... we had this thing (calories) always in our heads. It was terrible! We use to go to the gym...We did not even know what we were doing!
(Participant 1, Group 4)

In the words of these girls, feeling fat or thin is not something fixed by any objective standard. It depends on the perceptions and understandings that only make sense in their own environment. Their understanding guides their expectations and frustrations about their body weight. Sometimes the adolescents do not even realise why they are going on a diet– as reflected in the words of one of the girls “we did not even know what we were doing”. But the necessity to ‘feel good’ and accepted – which seems to be related to ‘feeling thin’ – drives their behaviour.

Category 4: Fitting in and not fitting in

The need to be socially accepted leads to our identification of a new category, which we call ‘Fitting in clothes’ and ‘Not fitting in clothes’, and it is exemplified in the following speech marks:

... You have to be thin to fit into the clothes that are in fashion... You have to look pretty in these clothes and not in the big ones. Well... I don’t know... I think it is because of appearance... to have a good looking body and be able to wear a two-piece bikini and not a one-piece swimsuit, you know? (Interview 1)

... The number 4 did not fit! (The participant is talking about a pair of jeans that she was trying on in a store) Can you believe it?! It did not fit! I could not fasten the button! I was so upset...so upset... I felt humiliated... I can’t even explain how I felt...I was already feeling a bit sad about myself, and then I had to say to the attendant in the store: “I did not fit in this one”... It was terrible... I knew then that I had to do something about my weight (Participant 1)... In our day we have to be the ones to fit in the clothes that are considered in fashion! It should be the opposite: the clothes should fit on us! And not that we should be trying to fit in the clothes! Why do we have to submit ourselves to that? (Participant 4, Group 4)

In the logic of these adolescents, ‘not fitting into the clothes’ means ‘feeling fat’. What is important is that they are not talking about a large range of sizes; they are talking about specific small sizes. In their understanding, to be and to feel thin, and consequently to feel socially accepted, means to ‘fit into small clothes’. There is no flexibility. The requirements to feel accepted and therefore to be comfortable with their bodies are strict and rigid.

In the discourse we can observe that the categories ‘fitting into clothes’ and ‘not fitting into clothes’ are not only related to clothes. The adolescents are also talking about feeling socially accepted or included. In the words of the girls, the wish to ‘fit into small clothes’ represents their wish to ‘fit into their peer group’. In other words, the concern about being overweight takes on social implications.

It is at this point that we introduce the final theme in our analysis, namely what we call ‘Exclusion and Inclusion’. This transition is given in Figure 5.

Theme 3: Exclusion and Inclusion

The theme ‘Exclusion and Inclusion’ emerges in the discourse out of the adolescent’s understandings about body weight and its link with their social relationships.

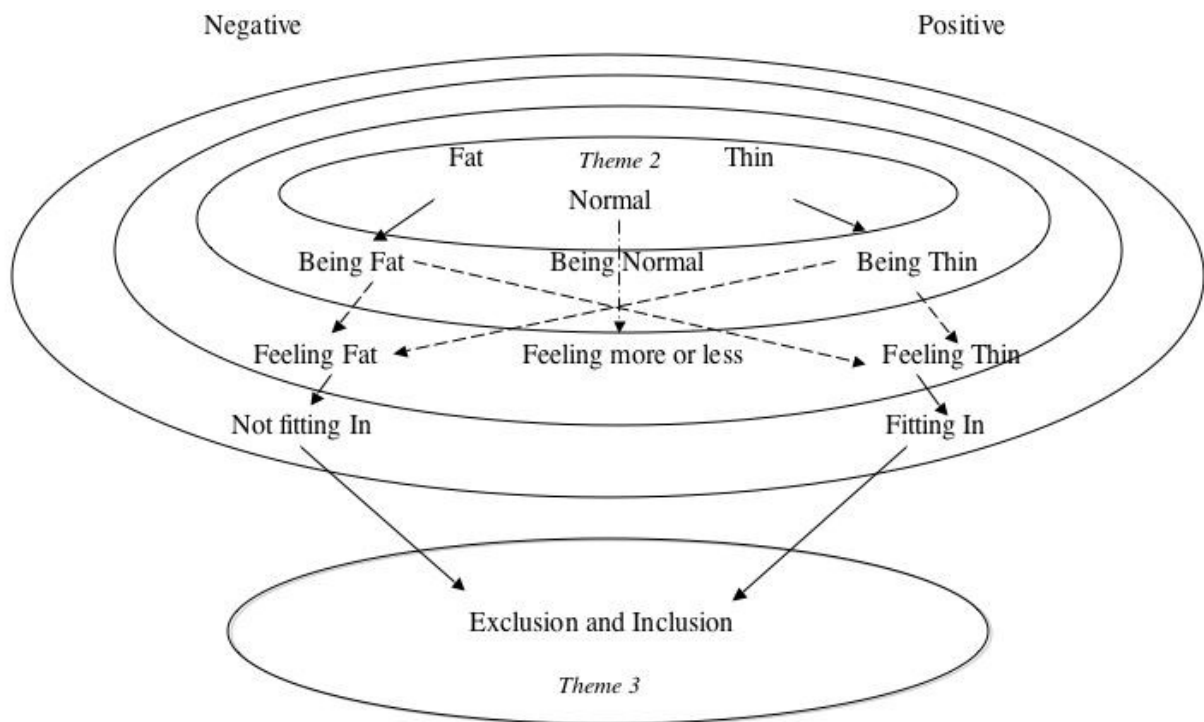


Figure 5. The emergence of the new theme, Exclusion and Inclusion

Physical appearance is seen as fundamental for establishing social relationships and to achieve acceptance within their peer groups. In the adolescents' social environment, the groups are determined by rigid criteria related to body standards. In other words, from their perspective the body condition determines 'who' will be excluded and 'who' will be included in a particular group.

...The skinny ones..., they are always surrounded by friends... In my school we have groups divided like this: the popular ones, that everyone likes, that are considered the good-looking ones... Then you have the nerds, those people that are nerds, you know? Then you have the normal ones, the ones in the middle, they are nothing... When a girl is skinny, and she is new at the school, for example, everyone is going to give her attention because she is thin. They will think that she is cool because she is skinny. If she is fat she cannot be cool, because she is fat, you know? These preconceptions... Discrimination against the fat ones... The fat ones are usually nerds and the thin ones have more of a chance to be popular... I don't think this is fair... (Interview 1)

...People judge you by your looks... They are not wondering if you are a nice person or anything like that. No... They are looking at you and thinking: "Is she pretty?" "Is she skinny?" "Is she fat?" (Participant 2). Yes... (Reflecting). This is true... (Participant 1). They don't even talk to you first! (Participant 2, Group 2)

This differentiation, which is determined by 'looks', influences their social relations in many different areas. The adolescent girls give a lot of examples about how their body weight is perceived by others in different situations and how it can determine their future life.

If you are fat the guys will make fun of you. For example, if you are fat and you like a guy he will make fun of you and you are going to feel really bad (Participant 2). If you are thin you are everything... (Participant 6). But there are some guys that like fat girls... (Participant 3). Not in this school (Participant 2). I think it depends on the guy (Participant 4, Group 1).

... If you think of a fat man and a thin woman..., and the other way around also... No one is going to like someone fat, you know? ...There is a lot of discrimination (Participant 1, Group 2).

...I always sang in a band... since I was little... My teacher says that I have a good potential to be a singer, but you know what worries me? To become fat in the future...Do you think that people would like my music even if I'm fat? ... I'll give you an example. If you have two singers, one is fat and the other one is skinny. Which one do you think would be more successful? The skinny one! Of course! Even if both sing well...the one that is going to be successful is the thin one, because everyone is going to think that she is beautiful. So then you end up with these kinds of anxieties about your professional life (Participant 1). Yes... there are some kinds of jobs that are impossible for a fat person, for example, a flight attendant (Participant 3, Group 2).

Looking closely at the above speech marks, we can clearly see that, in the understanding of these adolescents, body weight is a criterion for feeling empowered, acceptable and successful in their social relations. For them, being overweight is not only a problem of health or physical appearance. It is a problem also to the extent that it influences their social relationships, their perception of self, their self-esteem, and their understanding of what it means to be socially accepted and not excluded.

On the one hand, being overweight is linked to the idea of being excluded from different kinds of social relationships and a range of everyday situations. Sometimes it is not even a matter of having excess weight, but 'not fitting into' some patterns that are established in their world, like 'not fitting into small pants' that we pointed out before. On the other hand, being thin is seen as an 'open door' to a wide variety of opportunities. Put simply, to be thin means to be superior.

Category 1: Superiority and inferiority

The category which emerges from the theme 'Exclusion and Inclusion', in the adolescents' discourse, is that of 'Superiority and Inferiority'. This category represents what the adolescents see as a social consequence of the benefits and costs of the body condition. If being thin is a guarantee of social inclusion, then it follows that these adolescents will behave in a way which is conducive to weight control. If they are not overweight, they are afraid "to become fat in the future", as we saw earlier in the words of one of the participants.

However, what kind of behavioural consequence could we expect from this fear of becoming overweight? What is the consequence of their understanding of being fat as a criterion for exclusion and slenderness as a criterion for inclusion?

One consequence that comes from this fear of becoming fat is trying to lose weight.

Above we referred to quotes which illustrated some of the practices that the adolescents engaged in as a response to this fear. However, if the girls are able to maintain a slender weight, or if they succeed in losing weight, in their culture they have the right to feel superior.

...It is kind of normal that after you lose some weight, you feel full of yourself. You think: "Oh! Now I look good! Now I am pretty!"(Participant 2). After my cousin lost some weight she started..., I am not sure but..., she started to brag to the others, especially to her female friends. I think this is because..., when she was fat no one ever looked at her. She used to dress really bad before losing weight. (Participant 1, Group 2)

... The skinny girls tease the fat girls. They say: "I need to lose some weight", just to put down the fat girls. The fat ones don't argue...They keep quiet... We always

compare ourselves to each other and the thin ones can be really bad... I am like that sometimes. (Interview 2)

The discussion of this final category completes our analysis of the interviews and the ways in which these adolescents constructed their social representations of being fat or being thin. What now remains is to sum up the themes and categories which form a part of the integrated social representations.

TOWARD SOCIAL REPRESENTATIONS OF THE BODY

The social representations that emerge from this study extend across the boundary between the body and society. In effect, the social representations about the body which are constructed during the discourse affects the way these Brazilian female adolescents see themselves as fitting into society. A visual portrayal of the integrated social representations constructed by these Brazilian adolescent females is presented in Figure 6.

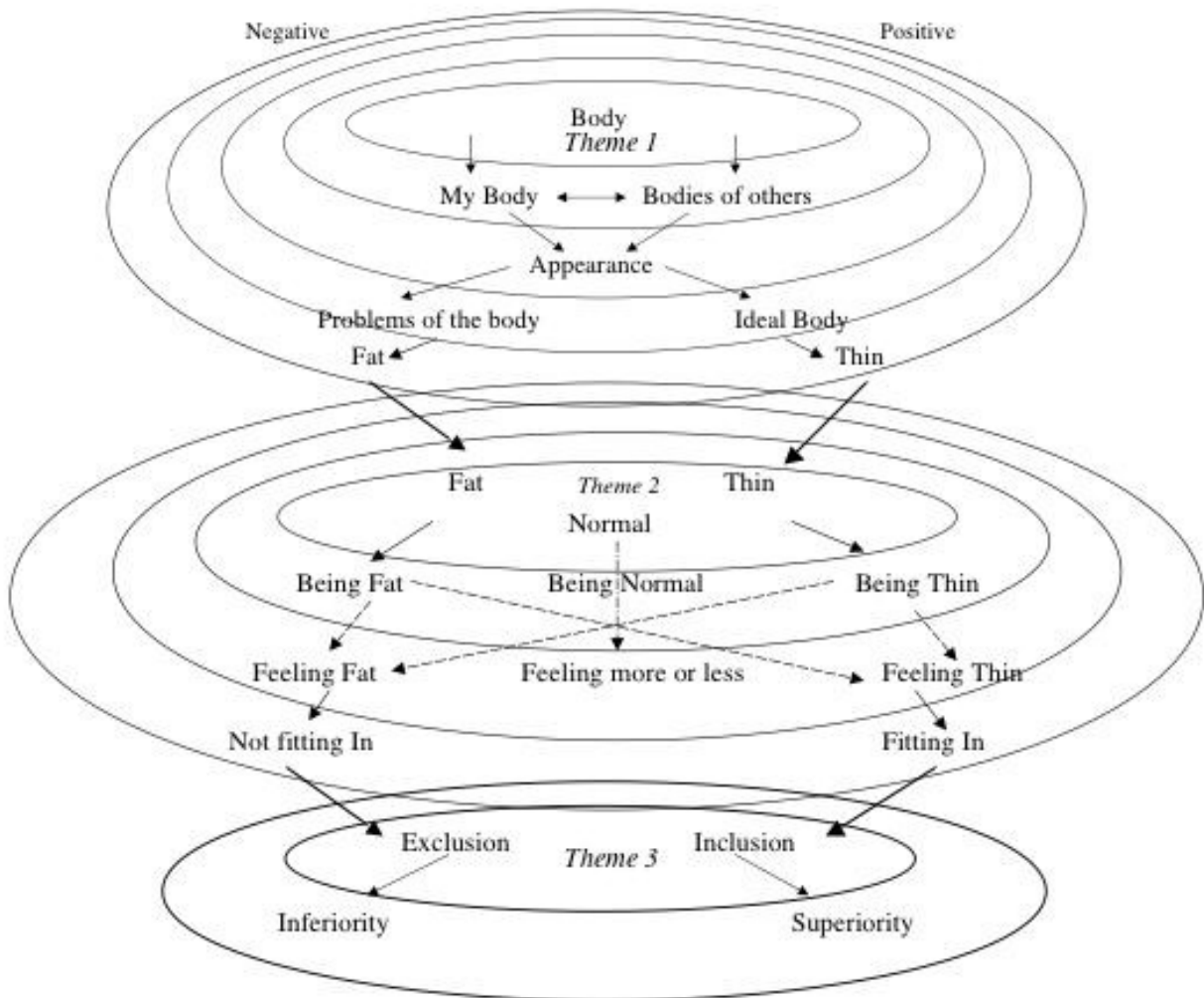


Figure 6. Social representations of the body among adolescent females in Brazil

The social representations of the body for these adolescents can be described in the following manner. First, the representations of being fat or thin emerge from the concept of the body. For these adolescents, the body represents positive and negative values, from which they focus on their own and other people’s bodily appearance. They tend to associate problems of the body with their own body, and the ideal body with that of others. The main characteristic of the body which differentiates problems from the ideal is that of being fat or being thin, which we call ‘body

weight'. The meanings of the body become, in the discourse, 'meanings of body weight', showing that 'body weight', for these girls, starts to turn into a criterion by which the adolescents measure their bodies and later their personal and social worth.

One of the first representations of body weight which emerges was related to the logic of these adolescents, whereby being fat and thin, and feeling fat and thin, are not necessarily related. Thus it is possible to be thin but to feel fat. Furthermore, in this logic there is no 'normal' condition. To be normal is to be 'nothing'. Thus, in the representations of these adolescents, the options are limited, and indeed the only choice is to be thin. And the choice becomes even more restricted when, according to their logic, a person can be thin but still feel fat. The final representation, and perhaps the most significant, is the fact that the adolescents see body weight as affecting their social relationships. Starting with the notion of 'fitting into clothes', their discourse leads to the notion of fitting into society. Thus the social representations of the body include the extent to which these adolescents feel that they 'fit' into their social environment, which includes social relationships, jobs, and other opportunities. Not to be thin means to feel excluded from this social environment. Finally, to the extent that a person feels thin and socially included, that person sees themselves as superior with a positive sense of self and high self-esteem.

From the above, we argue that the social representations of the body are strongly linked to body weight, and are also a powerful pressure on these young Brazilian female adolescents. To understand the complexity of these social representations helps to understand a range of behaviours which characterise youth, and in particular, behaviours related to weight problems and eating disorders.

PRACTICAL IMPLICATIONS OF SOCIAL REPRESENTATIONS OF THE BODY

The study of the social representations of the body has repercussions in different academic fields, from developments in social representation theory, to the field of overweight, obesity and eating disorders, and finally to studies of youth and education.

Social Representation Theory: Content and Process

The social representations that emerge from this study extend across the boundary between the body and society. The social representations of the body are built around social values that are not necessarily linked to body issues. However, asymmetric social relations such as exclusion, competition between peers, and feelings of superiority are understood by these adolescents as linked to body conditions. This shows that topics concerning weight problems, such as obesity and eating disorders, cannot be approached only as individual and medical conditions. The adolescents in this study show that it is not only the physical body which is important for understanding weight issues, but also the ways they perceive and represent their bodies in a social context that have an important place in their social practices.

Our study makes a further contribution to social representation theory at the methodological level. We have developed and used a systematic and rigorous procedure for the analysis of the *process* of the construction of social representations. We have respected the order of the appearance in the discourse of various stages of building their representations of being fat or being thin. Most of the previous studies have tended to analyse the content without showing the order and the way the content appeared in the process of construction. By building our analysis around the three operational concepts of themes, categories and units of analysis, as we explained at the outset of this paper, we argue that we have a better understanding of the content of the social representations, and the process by which they were constructed. Furthermore we have portrayed our analysis in a series of diagrams which visually describe both the content and the process. We

have incorporated in these diagrams both the contradictions inherent in social representation construction and the dynamism of the process itself.

Overweight, Obesity and Eating Disorders

With respect to the field of obesity and eating disorders, we are not passing judgement on being fat or thin, nor are we trying to resolve the social problem of obesity. Instead we are investigating the understandings of these conditions by these groups of adolescents, and how these understandings lead to particular types of behaviours. Most of the existing studies on the vulnerability of adolescents to weight problems and eating disorders show that they are a group at risk because they have a high probability to go on diets even when they do not need to do so. For most adolescent girls, the 'fear of becoming fat', an emotion understood as one of the symptoms of anorexia and bulimia, is becoming more common than researchers have expected. Recent studies, such as that conducted by Nunes and her colleagues (Nunes, et al., 2000; Nunes, et al., 2003) suggest that the fear of becoming fat might be a shared anxiety in adolescent culture generally.

Our study makes a further contribution to the above, and shows 'why' adolescents do this. They fear being fat because being fat means social exclusion. What appears to be logical for the adolescents does not appear to be logical from the perspective of outsiders. If you are an adolescent female who fears getting fat, and you know the social consequences of getting fat, then even if you are thin, you go on a diet. The social representation of feeling fat or feeling thin does not have to correspond with the objective condition of being fat or being thin in order to have an impact on the weight-related behaviour of these adolescents. This suggests that what really matters is not reality itself, but the way the reality is represented and understood in the minds of these individuals. For them, the social representation becomes the reality.

Therefore to be on a diet is the only way for these adolescent girls, in their understanding, to guarantee being thin, and to be thin means to be automatically included in social relationships, and to be socially superior. After all, even a person who is thin and still goes on a diet, experiences higher self esteem and perceives their admiration in the eyes of others.

In the logic of these adolescents, to be 'normal' is not relevant or interesting. The focus here is more on being socially included and not excluded. Their perception of 'normal weight' does not seem to guarantee that they will be automatically included in the group, whereas being slender does. Therefore their social representations of body weight do not give any alternative except to concentrate on keeping their weight down, and this means an obsession with dieting. In extreme cases, this obsession might develop into other more serious eating disorders such as anorexia or bulimia.

Studies of Youth

With respect to the studies of youth, this study contributes to the understanding of the complex process involving puberty physical changes and their social psychological consequences. In addition, it is not simply the physical changes themselves that are of concern, but the impact that the social representations of these changes has for the adolescents in the study.

This study also confirms already existing research findings that show that the perception of the body for female adolescents is strongly linked to appearance and weight issues (Coleman and Hendry, 1999). However, our study takes a further step. We show that the content of their perceptions of their own bodies are not only related to weight issues, but are also related to social issues, both with a strong negative bias. Furthermore, we show that the process of the construction of these perceptions is a part of shared understandings (social representations) and not the product of the aggregation of individual constructions.

The importance of these social representations is inflated by the fact that, at this stage in their physical and social development, young females have little in the way of social accomplishment with which to evaluate themselves, and with which to compare themselves to others. Thus their body form and overall physical appearance take on significance beyond their physical concerns, and extends into their social concerns as well. This is reflected in the social representations of the body constructed by our participants, and it explains why these representations eventually included the theme of social inclusion and exclusion. Future research on the body images of adolescent females, and males as well, might focus more closely on the unique role that the body plays in self evaluation and social comparison, given that at this age they have had few opportunities for other accomplishments or achievements against which to evaluate and compare themselves.

Implications for Schools and Educational Programs

Given that the pressures for these representations of the body come from the media, parents, peer groups and the school, further studies are needed about possible intervention strategies which can modify the construction process to prevent the extreme negative and potentially destructive content of these representations.

Research on the influence of the media has been extensive (Botta, 1999; Herbozo, Tantleff-Dunn, Gokee-Larose, and Thompson, 2004; Posavac and Posavac, 2002), and that of parents and peer groups appears to exist, but to a lesser extent (Phares, Steinberg, and Thompson, 2004). While it may be desirable that intervention strategies be developed regarding the media, parents and peer groups, in reality such strategies are impractical and not likely to occur. However, intervention at the school level is a different matter.

The students in this study represented levels of schooling from high school to university, and yet the representations of the body were consistent. Thus, it has been argued that the best type of intervention for youth is 'psychoeducational', that is, the inculcation of media literacy skills in order to have a more critical approach to the messages from the media and other agents (Phares et al., 2004; Tiggemann, 2004). In addition, it has been suggested that prejudices in attitudes and practice toward certain body forms in school settings need to be removed (Greenleaf and Weiller, 2005). Both of these suggested strategies imply the development of intervention programs in schools which will empower young people to be able to interpret critically, and even resist the negative influence of the media and other similar agents in their day-to-day construction of social representations about the body.

School-based programs might also introduce strategies, which will provide a broader base for adolescents to evaluate and compare their own self worth. Other aspects of the body beside body weight and body form might be emphasised, such as eyes or hair. Positive aspects of personality characteristics and academic and other achievements, such as sport or other manual skills need to be stressed. In this way, body form and body image might be put in a more broad and balanced perspective, alongside other positive features of the total person.

Limitations of the Study

Some might argue that our analysis is based on a relatively small number of focus groups and participants. However, as we noted, the number of our focus groups is consistent with the recommendations of Morgan (1997) and Bloor and his colleagues (2001), and is sufficient to attain 'saturation'. Furthermore, we believe that the methodological procedures that we used, and the depth of the focus interviews themselves, offset any concerns regarding the numbers of groups and participants.

A common criticism of the use of focus groups is that the participants sometimes reveal what they think the researcher wants to hear. In this study, this concern was minimised by the way the focus group interviews were conducted. As pointed out, the interviewer did not intervene in the discussion until the key topic of body weight was mentioned. In this way it was the participants who guided the direction of the discourse and not the interviewer. (See Morgan, 1997.)

A final concern with social representation research is how to account for the social representations of the researchers themselves, and how they might contaminate the analysis of the discourse. In response, we argue that the systematic and rigorous methodology that we used served as a barrier and protector between any prior representations of the researchers and those being constructed by the participants.

In effect, the limitations that we have identified above are not unique to social representation research, and are common to all research of this nature. However, we have indicated our recognition of these limitations and how we think we have overcome them.

FUTURE RESEARCH

Future research needs to expand the scope of the present study in a number of ways. First, this study needs to be replicated with other groups of Brazilian females, especially in other regions and perhaps with a wider age range. Second, future research needs to replicate this methodology with other groups of females in a cross-cultural context in order to determine whether the same social representations will be constructed. Third, the study should be conducted on males to determine whether their own constructed representations are similar or different from those of females. Finally, it would be possible to approach the study of weight-related social representations from other theoretical perspectives, for example the ideological perspective, which would include an emphasis on the role of external domination agents such as the media and the fashion industry.

By pursuing the study of weight-related social representations using different groups and different theoretical perspectives, not only will our knowledge of the process of the social construction of social representations be advanced, but also our knowledge of the content regarding body issues will be broadened. There are both academic and policy-related benefits for continuing this research agenda.

The functions of the peer group, which emphasises the messages conveyed by the media, particularly through television advertising, also warrant further investigation. The increasing levels of retention at school for 12 or more years, and the growing levels of participation in universities and colleges at a time in the lives of young people when they become sexually aware and active, suggest that social representations of the body, as described in this article, takes place in the classrooms and on the playgrounds of educational institutions where the youth of today are concentrated. Consequently, not only do the various educational institutions have an opportunity to develop among youth a critical approach to these media messages, but also it seems likely that the issues raised in this article could be a hidden factor in explaining differential success at school and continued participation in education.

As educational provision and participation expands in developing countries, the body image issues addressed in this article move beyond Western societies. These emergent critical issues regarding body image affect the growth of health and a sense of well-being among youth. Furthermore, as was clear in our findings in this Brazilian study, they are perceived to affect future democratic involvement in community life and economic participation in the workforce. Clearly, the study of social representations of the body among youth, and their implications for education, is an important subject for continued research in all societies and cultures.

REFERENCES

- Bauer, M. W. and Gaskell, G. (1999). Towards a Paradigm for Research on Social Representations. *Journal for the Theory of Social Behaviour* 29, 163-186.
- Bloor, M., Frankland, J., Thomas, M., and Robson, K. (2001). *Focus Groups in Social Research*. London: Sage Publications.
- Botta, R. A. (1999). Television Images and Adolescent Girls' Body Image Disturbance. *Journal of Communication* (Spring), 22-41.
- Brush, Hilde (1973). *Eating Disorders: Obesity, Anorexia Nervosa and the Person Within*. New York: Basic Books Publishers.
- Chernin, K. (1981). *The obsession: Reflections on the Tyranny of Slenderness*. New York: Harper and Row.
- Coleman, J. C. and Hendry, L. B. (1999). *The Nature of Adolescence*. London: Routledge.
- Cash, T. F. (2004). Body image: Past, present, and future. *Body Image*, 1, 1-5.
- Duveen, G. (2000). *Some Problems in the Diffusion of Social Representations*. 6th International Summer School on Social Representations, Rome.
- Fallon, P., Katzman, M.A. and Wooley, S. C. (1994). *Feminist Perspectives on Eating Disorders*. The Guilford Press: New York.
- Frost, L. (2001). *Young Women and the Body: A Feminist Sociology*. New York: Palgrave.
- Garine, I. de and Pollock, N. J. (Eds.) (1995). *Social Aspects of Obesity*. Amsterdam: Overseas Publishers Associates.
- Germov, J. and Williams, L. (1999). Dieting Women: Self-Surveillance and the Body Panopticon. In J. Sobal and D. Maurer (Eds.), *Weighty Issues: Fatness and Thinness as Social Problems* (pp.117-132). Aldine de Gruyter: New York.
- Gordon, R. (1990). *Anorexia and Bulimia: Anatomy of a Social Epidemic*. Cambridge, MA: Basil Blackwell.
- Greenleaf, C., and Weiller, K. (2005). Perceptions of youth obesity among physical educators. *Social Psychology of Education*, 8(4), 407-423.
- Haworth-Hoeppner, S. (1999). Medical discourse on Body Image: Reconceptualizing the differences between women with and without eating disorders. In J. Sobal and D. Maurer (Eds.), *Interpreting Weight: The Social Management of Fatness and Thinness* (pp.89-111). New York: Aldine de Gruyter,.
- Herbozo, S., Tantleff-Dunn, S., Gokee-Larose, J., and Thompson, J. K. (2004). Beauty and Thinness Messages in Children's Media: A Content Analysis. *Eating Disorders*, 12, 21-34.
- McKinlay, A., and Potter, J. (1987). Social Representations: A conceptual critique. *Journal for the Theory of Social Behaviour*, 17, 471-487.
- Morgan, D. (1997). *Focus groups as Qualitative Research*. Thousand Oaks: Sage Publications.
- Moscovici, S. (1976). *La psychanalyse: Son Image et son Public* (2ème edition). France: Presses Universitaires de France. (1st Edn. original 1961).
- Moscovici, S. (1987). Answers and Questions. *Journal for the Theory of Social Behaviour*, 17, 513-529.
- Moscovici, S. (1988). Notes Towards a Description of Social Representations. *European Journal of Social Psychology* 18, 211-250.
- Nasser, M. (1997) *Culture and Weight Consciousness*. New York: Routledge.
- Nunes, M. A., Barros, F. C., Olinto, M. T. A., Camey, S., and Mari, J. de J. (2003). Prevalence of Abnormal Eating Behaviors and Inappropriate Methods for Weight Control in Young Women from Brazil: A Population-Based. *Eating and Weight Disorders* 8, 100-106.

- Nunes, M. A., Olinto, M. T. A., Barros, F. C. and Camey, S. (2000). A Influencia da Percepcao do Peso Corporale e do Indice de Massa Corporal nos Comportamentos Alimentares Anormais em Mulheres Jovens de Porto Allegre, RS (Influence of Body Mass Index and the Perception of Body Weight on Abnormal Eating Behavior in Young Women From Porto Alegre, RS). *Revista Brasileira de Psiquietrie* 23, 21-27.
- Orbach, S. (1978). *Fat is a Feminist Issue: A Self-help Guide for Compulsive Eaters*. New York: Berkeley Book.
- Paquette, M-C. and Raine, K. (2004). Sociocultural context of womens' body image. *Social Science and Medicine*, 59, 1047-1058.
- Phares, V., Steinberg, A. R., and Thompson, J. K. (2004). Gender Differences in Peer and Parental Influences: Body Image Disturbance, Self-Worth, and Psychological Functioning in Preadolescent Children. *Journal of Youth and Adolescence*, 33, 421-429.
- Posavac, S. S. and H. D. Posavac (2002). Predictors of Women's Concern with Body Weight: The Role of Perceived Self-Media Ideal Discrepancies and Self-Esteem. *Eating Disorders*, 10, 153-160.
- Semin, G. R. (1995). Social Representations. In A. S. R. Manstead and M. Hewstone (Eds.), *The Blackwell Encyclopedia of Social Psychology* (pp. 601-605). Oxford: Blackwell Publishers.
- Sobel, J. and D. Maurer, (Eds.) (1999). *Interpreting Weight: The Social Management of Fatness and Thinness*. New York: Aldine de Gruyter.
- Spink, M. J. (1995). Desvendando as teorias implícitas: uma metodologia de análise das representações sociais. In P. A. Guareschi and S. Jovchelovitch (Eds.), *Textos em representações sociais*. 2nd ed. Petropolis, R.J.: Vozes.
- Stewart, D.W. and Shamdasani, p.N., (1990) *Focus Groups: Theory and Practice*. Newbury Park: Sage Publications.
- Stearns, p.N. (1997) *Fat History: Bodies and Beauty in the Modern West*. New York: New York University Press.
- Stinson, K. (2001). *Women and Dieting Culture: Inside a Commercial Weight Loss Group*. New Jersey: Rutgers University Press.
- Székely, É. (1988). *Never too Thin*. The Women's Press: Toronto, Canadá.
- Tiggemann, M. (2004). Media Influences on Body Image Development. In T. F. Cash and T. Priuzinsky (Eds.), *Body Image: A Handbook of Theory, Research, and Clinical Practice* (pp. 91-98). New York: Guilford Press.
- Wolf, Naomi. (1990). *The Beauty Myth*. London: Chato and Windus.

Using regression analysis to establish the relationship between home environment and reading achievement: A case of Zimbabwe

Gibbs Y. Kanyongo

School of Education, Duquesne University *kanyongog@duq.edu*

Janine Certo

School of Education, Duquesne University *certo@duq.edu*

Brown I. Launcelot

School of Education, Duquesne University *brownli@duq.edu*

In this study, we report results of a study examining the relationship between home environment factors and reading achievement in Zimbabwe. The study utilised data collected by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). The data were submitted to linear regression analysis through structural equation modelling using AMOS 4.0. In our results, we showed that a proxy for SES was the strongest predictor of reading achievement.

Zimbabwe, reading achievement, home environment, linear regression,
structural equation modelling

INTRODUCTION

Past research has indicated that a significant relationship exists between children's home environment and reading achievement. However, most such studies have been conducted in Western countries where the concept of home environment is different from that in developing countries. In developed countries, almost all students have amenities like electricity and piped water in their homes, and these factors are never thought of being influential in a student's academic performance. In the current study, the home environment factors considered among other factors were: possession of such things like piped water, electricity, refrigerator and TV. Collectively, these possessions were taken to be a measure of social economic status (SES). SES, together with several other factors were used as measures of the students' home environment. The student's score on a reading test was used as a measure of reading achievement. The study utilised a sample of 2697 sixth grade students who were randomly sampled across the schools in Zimbabwe.

PURPOSE OF THE STUDY

The purpose of this study was to investigate the relationship among home environment variables and reading scores among Grade 6 students in Zimbabwe. The study is based on the data collected by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) during the period 1995 – 1998 under the auspices of the International Institute for Educational Planning (UNESCO-IIEP, 2004). The data were collected during the first educational research policy project named SACMEQ I covering seven Southern African countries, and Zimbabwe was one of the seven countries. Specifically, the current study seeks to answer the following two major questions:

- a) What family environment variables are predictive of reading achievement among Grade 6 students in Zimbabwe?
- b) How strong are family environment variables at predicting reading scores among Grade 6 students in Zimbabwe?

PREVIOUS RESEARCH AND THEORETICAL FRAMEWORK

There is a considerable body of theoretical and empirical work that suggests that home environment, in general, plays a crucial role in student learning and achievement (Walberg, 1999). Research focusing on home environment variables continues to be of merit, for Parcel and Dufur (2001), in their National Longitudinal Survey of Youth (NLSY), found that parental and material resources in the home (what they termed 'capital effects') were stronger than school effects.

The theoretical mechanisms that may explain the effect of home environment on child outcomes are related and numerous. Social capital theory (Coleman, 1988; Parcel and Dufur, 2001). Thus, social capital and resources at the country or family level, provide students with advantages for being successful at school (Aru, 1998; Bradley and Corwyn, 2002; Heyneman and Loxley, 1982). Such resources also benefit students indirectly through an increased number of cultural opportunities and better health standards (Murphy et al., 1998; Neisser et al., 1996).

A causal mechanism discussed in the literature is socio-economic status (SES). It is no longer questioned that low income leads to negative consequences for children (Duncan and Brooks-Gunn, 1997). Children living in poverty face developmental deficits that are most likely due to the inability of high-poverty families to provide adequate food, shelter, and other material goods that foster the healthy cognitive development of children (Duncan, Brooks-Gunn, and Klebanov, 1994; Hanson, McLanahan, and Thomson, 1997; Korenman, Miller, and Sjaastad, 1995). A meta-analysis of 101 related studies (White, 1982) suggested that when home atmosphere or environment was used as an indicator of SES, relatively high correlations were found between SES and academic achievement.

Chui and Khoo (2005) found that 15-year-old students across 41 countries (N = 193,076) scored higher on tests in mathematics, reading, and science when they had more economic resources in their country, family and school. The positive log-linear effect of per capita GDP was consistent with past research showing that students in richer countries benefited from more nutritious food, books in the home, and better health care, all of which, in turn, supported higher academic performance (Alaimo, Olson, and Frongillo, 2001; Murphy et al., 1998; Neisser et al., 1996). Results were consistent with other studies in that girls scored higher in reading but lower in mathematics and science (National Center for Education Statistics, 2003; Third International Mathematics and Science Study Group, 1995).

In general, parent job status and parent education seems to have different effects (Chui and Khoo, 2005). Parents' higher job status and mother's education improved the academic performance of their child. A father's level of schooling, however, did not. These results mirror other studies showing that parents' social networks and mother's schooling affected students' academic performance more than did father's level of schooling (Stafford and Dainton, 1995).

Researchers have not only found that adolescents from lower income and less-educated families performed less well in school, but those from single-parent and large families fared less well in school than their counterparts from higher income, better-educated, two-parent and small families (Astone and McLanahan, 1991; Sputa and Paulson, 1995). Stevenson and Baker (1987) found that the educational level of parents explained more of the variability in school achievement than did other family demographic characteristics. This might be explained by the theory that caregivers with more 'human capital' had greater education and skills, which they could draw upon to teach

their children cognitive and social skills and social and cultural norms (Och, Taylor, Rudolph, and Smith, 1992; Snow, Perlmann, Berko Gleason, and Hooshyar, 1990; Swick and Broadway, 1997). Again, since mothers were the primary caregivers in many homes, their human capital generally affected their children's performances more than did the a father (Darling-Fisher and Tiedje, 1990; Emery and Tuer, 1993; Stafford and Dainton, 1995).

There are factors in the home environment that have all been found to be associated specifically with reading achievement. Not surprisingly, these include financial and material resources (Grissmer et al., 1994; Parcel and Menghan, 1990, 1994a, 1994b; Saracho, 1997a), number of children in the home (Blake, 1989; Coleman, 1988, 1990; Downey, 1995; Parcel and Menghan, 1990, 1994a, 1994b), parent education (Grissmer et al., 1994; Saracho, 1997a), mother's age (Grissmer et al., 1994), and caregiver investment in age-appropriate cognitive stimulation (Leseman and de Jong, 1998; White, 1982). The data are mixed, however, in regards to which home environment variables best predict reading success.

Several studies have demonstrated that increased numbers of children within the family lead to less favorable child outcomes, in general, probably due to resource diffusion (Blake, 1989; Downey, 1995; Parcel and Menaghan, 1990, 1994a, 1994b). Parcel and Menaghan (1993) found that children from larger families had lower levels of reading achievement. Coleman (1988, 1990) argued that even if the bond between caregiver and child was strong, it would not automatically result in improved child outcomes unless the child had access to parental resources.

Researchers at the Rand Institute on Education and Training (Grissmer et al., 1994) found a relationship between children's reading achievement and parent's level of education, family income, and the mother's age. Although parent's education, occupation, and income have been found to be related to children's reading outcomes, some suggest that the actual characteristics of the home literacy environment that was created by the caregiver might be more important (Leseman and de Jong, 1998; White, 1982).

The recent term 'home literacy environment' has generally referred to participation in literacy-related activities relegated to availability of print material and frequency of reading (Leseman and de Jong, 1998). Saracho (1997a) proposed that parent's literacy level and the availability of reading materials worked together as the primary characteristics of the home environment that related to a child's literacy development. Other research has extended the scope of the home literacy environment (Christian, Morrison, and Bryant, 1998; Griffin and Morrison, 1997; Payne, Whitehurst, and Angell, 1994) to include such variables as the age of the child when joint reading began, independent child or caregiver reading, and frequency of behaviours that interfered with reading (for example, television viewing).

A factor often overlooked in research on the home environment is that a child may not live with a parent at all or may live away from them during the school week. Stress theory holds that changes in family organisation or circumstances might cause stress in children's lives (Elder, 1974). Family events may directly increase children's stress because of changes in household composition or changes in residential location. In response to these changes, children would be alienated from or might disengage from the home environment and receive less parental nurturing and socialisation.

Greenstein originally (1954) noted that children with television in the home had higher grade point averages than their counterparts without television in the home, a phenomenon most likely related to socio-economic status. More recent research addressing the relationship between television viewing and achievement has yielded contradictory findings. For example in small-scale studies controlling for IQ, SES and parental education, no relationship has been found between viewing and achievement (Ritchie, Price, and Roberts, 1997). However, secondary analyses using large-scale data sets have demonstrated that for adolescents, television viewing

time is negatively related to achievement, particularly with regard to reading performance (Neuman and Prowda, 1982; Williams, Haertel, Haertel, and Walberg, 1982).

Home environment and reading achievement research has been largely dominated by a focus on early reading acquisition, while research on the relationship between home environments and reading success with preadolescents (Grades 4-6) has been largely overlooked. There are other limitations as well. Clarke and Kurtz-Costes (1997) argued that prior research has failed to examine the relationship between certain home environment variables and achievement of children of poverty. For example, if having a television (and television viewing) does have positive influences on children's developing academic skills, one might expect similar effects to be found in economically-disadvantaged homes where children typically have fewer educational resources available to them, such as books and other literacy materials.

Because of the large numbers of students living in poverty in the world, more data are needed to establish an understanding of the relationship between home environment and reading success, particularly for preadolescent students. Scholars have argued further that researchers should limit their conceptualisation of middle-class White Americans as the norm and should view other ethnic and social class groups only in terms of their deviance from this norm (Ogbu, 1981; Scott-Jones, 1984, 1987; Slaughter and Epps, 1987). This research seeks to examine an economically disadvantaged group as a distinct entity.

METHOD

Participants

This study utilised data collected by SACMEQ during an educational research policy project. The data consisted of 2697 sixth grade students (1329 boys and 1368 girls) who were randomly selected across Zimbabwe as part of a study undertaken during the period 1995 to 1998 to collect data on the conditions of schooling and the quality of education in seven Southern African countries. The mean age of the participants at the time of data collection was about 12 years.

Treatment of the data

The original data obtained from SACMEQ were already treated for missing values and screened for outliers. In the current study, the descriptive statistics for all the variables were examined to make sure they fell within acceptable range and skewness is one such statistic that was carefully looked at. Histograms were obtained for all the variables whose skewness statistic was greater than 1 to have a pictorial view of the distribution of the variables. Only one variable, STAY, was found to be highly skewed and its treatment is described in the section below.

Variables

The dependent variable in this study was **READING SCORE** and it was measured by a reading test. The independent variables were several home environment variables. These variables that were based on recommendations from previous research were selected for use in this study. Four home environment variables were combined to be a measure of SES. These variables are: (a) Possession of a TV, (b) Possession of a refrigerator, (c) Possession of piped water, and (d) Possession of electricity. Five other variables, (a) Someone makes sure you did homework, (b) Someone helps with homework, (c) Someone asks you to read to him or her, (d) Someone questions on what you read, and (e) Someone looks at school work were combined to form a variable, **HOME**.

The other variables considered in the present study are: **GENDER**, **READING AT HOME**, **STAY**, **MEALS**, and **BOOKS**.

The variable, *READING AT HOME*, measured whether students did any reading at home. The variable, *STAY* measured the place where the student stayed during school week and it was originally measured on a 4-point scale with 1 being 'stay with parents', 2 'stay with relatives', 3 'stay in a hostel', and 4 'stay by myself'. However, since the distribution was highly skewed, with more than two thirds choosing category 1, categories 2, 3, and 4 were combined, yielding two categories; 1 'with parents' and 2 'other'. The variable, *MEALS*, measured the number of times that the students ate breakfast, lunch and supper per week, and *BOOKS* measured the number of books that the students had at home. Information about the independent variables and how they were measured is provided in Table 1.

Table 1. Summary of the independent variables (student home environment variables) and how they were measured, along with sample means and standard deviation ($n = 2697$)

Variable	Scale	M	SD
Reading Score		25.60	10.05
Gender		1.51	0.50
Boys	1		
Girls	2		
Place the student stays during school week		1.25	0.43
With parents	1		
Other	2		
Reading at home		1.74	0.44
Never	1		
Sometimes	2		
SES (Combined the following)		1.24	0.35
<i>Possession of a TV</i>			
<i>Possession of a refrigerator</i>			
<i>Possession of piped water</i>			
<i>Possession of electricity</i>			
No	1		
Yes	2		
Meals (combined the following)		2.86	0.60
<i>Breakfast</i>			
<i>Lunch</i>			
<i>Supper</i>			
Not at all	1		
1 or 2 times per week	2		
3 or 4 days per week	3		
Everyday	4		
Home (combined the following)		1.26	0.44
<i>Someone makes sure you did homework</i>			
<i>Someone helps with homework</i>			
<i>Someone asks you to read to him/her</i>			
<i>Someone questions on what you read</i>			
<i>Someone looks at school work</i>			
Never	1		
Sometimes	2		
Most of the times	3		
Number of books at home		2.28	1.14

RESULTS

The results are presented according to the research questions asked in this study.

Research question 1: What home environment variables are predictive of reading achievement among Grade 6 students in Zimbabwe?

This question was answered by obtaining the regression weights and the p-values for these weights. In Table 2, the results of the regression analysis are displayed. In the table, the estimated regression weights, standard errors and p-values for all the predictors are given. *GENDER* was not a significant predictor of reading achievement, ($\beta = 0.03$, $p = 0.13$), so was *BOOKS* ($\beta = 0.02$,

$p = 0.39$). However, STAY ($\beta = -0.13$, $p < 0.00$); SES ($\beta = 0.27$, $p < 0.00$); MEALS ($\beta = 0.11$, $p < 0.00$); HOME ($\beta = 0.08$, $p < 0.00$), and READING AT HOME ($\beta = 0.17$, $p < 0.00$) were all significant predictors of reading achievement. The variable STAY had a negative relationship with READING SCORE. This makes sense since this variable was coded as 1- stay with parents and 2 – stay with others. Students who stayed away from parents tended to perform poorly academically.

Table 2. Regression Weights

	Estimate	S.E.	P
Reading score <-- Gender	0.52	0.34	0.13
Reading score <-- Stay	-2.90	0.40	0.00
Reading score <-- SES	7.71	0.52	0.00
Reading score <-- Books	0.15	0.17	0.39
Reading score <-- Meals	1.88	0.30	0.00
Reading score <-- Home	1.50	0.34	0.00
Reading score <-- Home reading	3.74	0.44	0.00

Research question 2: How good are family environment variables at predicting reading scores among Grade 6 students in Zimbabwe?

In order to answer this question, we obtained the standardised regression weights of all the variables. The standardised regression weights give the weighted contribution of each predictor to the dependent variable. Using the standardised weights makes sense since the variables are measured in different units. Another advantage of using the standardised regression weights is that these weights can be taken as measures of effect size. The effect sizes provide an indication of the practical importance of each predictor. The larger the standardised regression weight (effect size), is that variable in predicting the dependent variable.

Besides the standardised weights, we also obtained the squared multiple correlation. The squared multiple correlation provides a measure of the contribution of all the predictors taken together. These results are displayed in Table 3. SES was the strongest predictor as indicated by its estimated standardised regression weight, ($\beta = 0.27$, $p < 0.000$). The estimated standardised regression weights show the relative importance of each predictor in the model. Collectively, all the seven predictors account for about 21% of the variance in reading achievement ($R^2 = 0.21$).

Table 3. Standardised Regression Weights

	Estimate
Reading score <-- Gender	0.03
Reading score <-- Stay	-0.13
Reading score <-- SES	0.27
Reading score <-- Books	0.02
Reading score <-- Meals	0.11
Reading score <-- Home	0.08
Reading score <-- Home reading	0.17

Squared Multiple Correlations Reading score = 0.21

DISCUSSION

The home environment variables considered in this study were GENDER, STAY, READING AT HOME, BOOKS, MEALS, HOME, and SES. The results of these analyses demonstrate that, of these variables, SES – limited here to television, refrigerator, piped water, and electricity possession – was the strongest predictor of reading achievement. This is consistent with previous research that material resources in the home are associated with children's reading achievement (Grissmer et al., 1994; Parcel and Menghan, 1990, 1994a, 1994b; Sarracho, 1997a; White, 1982).

In short, these findings mirror other studies that in poorer countries, this decreased financial capital is disadvantageous to children's achievement (Aru, 1998; Bradley and Corwyn, 2002; Heyneman and Loxley, 1982). Previous research also reports that a lack of resources which could provide better health standards (such as piped water), can indirectly affect children's success at school (Duncan, Brooks-Gunn, and Klebanov, 1994; Hanson, McLanahan, and Thomson, 1997; Korenman, Miller, and Sjaastand, 1995; Murphy et al., 1998; Neisser et al., 1996).

The findings revealed a negative relationship between STAY and reading achievement. This is consistent with previous research that students living with parents tend to do better than those who live with people other than their parents. These results may give credence to stress theory (Elder, 1974), concerned with children who are physically removed from parental nurturing and socialisation. The SACMEQ reading study questionnaire did not include information on the number of siblings, so we were unable to confirm findings from previous research (Astone and McLanahan, 1991; Sputa and Paulson, 1995) showing that adolescents from smaller families fare better in reading.

Overall, the contribution of home environment factors considered in this study was about 21 per cent of the variance in reading achievement. That means that home environment factors account for a fifth of the variance in reading achievement. This is an adequate proportion considering the fact that from previous studies, the greatest predictor of reading achievement is prior reading achievement. Although this may seem obvious, it is important to note that most previous studies conducted on reading achievement have been done in Western settings. The implications of the current study can never be over-emphasised. Home environment plays a critical role in influencing reading achievement in schools and SES plays an even greater role in influencing reading achievement. These findings seem to indicate that it does not matter much how SES is measured in the different cultural contexts, the results show that SES is an important predictor of achievement. We know that in Western contexts, SES is measured differently from the way it is measured in this study, but it is still a good predictor of reading achievement.

Considering the fact that more and more students in Zimbabwe are becoming orphaned most family structures are being decimated by the AIDS pandemic, the impact of these findings become all too apparent. In this study, we have demonstrated the importance of having students stay with parents during school days. This is very important because parents provide the necessary monitoring that deters children from engaging in mischief and that probably channels their energy into academic pursuits. However, we are aware that this is not the whole story because in Zimbabwe, boarding schools (where students stay away from parents during school days) are some of the best schools in terms of pass rates. So on the surface, these results appear to be a contradiction with reality. However, we also know that there are many other children in Zimbabwean schools who live by themselves either because they are orphans or because they live in make-shift houses close to schools mainly because the schools are far away from their homes and their parents cannot afford boarding fees. These children have no supervision and it is only logical to assume that most of them spend very little time engaging in school work.

The contribution of certain individual predictors like BOOKS at home was surprisingly low, a finding inconsistent with previous research (Saracho, 1997a) and the general theme in the literature that increased educational resources is an important factor in student achievement (Alaimo, Olson, and Frongillo, 2001; Grissmer et al., 1994; Murphy et al., 1998; Neisser et al., 1996; Parcel and Menghan, 1990, 1994a, 1994b). One would think that the number of books at home would be a significant predictor of reading achievement. Probably part of the reason for the weak contribution of number of books is the way the data were collected. Students were not specifically asked for the number of 'reading books', but the number of books in general. So there is a possibility that students merely indicated any books they had at home even those they would never read. So we viewed this as a limitation of the study. Another limitation of the instrument is

that it was not developed to collect accurately data on the home literacy environment, a complex compilation of variables that would include the age of the child when joint reading began, independent child or caregiver reading, and the frequency of behaviours that interferes with reading (Christian, Morrison, and Bryant, 1998; Griffin and Morrison, 1997; Payne, Whitehurst, and Angell, 1994). Such data were not collected by the instrument used, nor was the full range of possible parent-child literacy-related activities considered.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The instrument used to collect the data used in this study was not specifically developed to collect data on home literacy. Furthermore, of those instruments that were available to measure home literacy environment, they were developed to do so mainly in Western cultures. The home literacy environment for developed nations is completely different from that in less developed nations. As such, developing a culturally appropriate instrument to measure home literacy environment in less developed nations would produce more reliable results. We see this as a possibility for future research. Also, it would be important to include the full range of possible parent-child literacy-related activities if such an instrument was to provide a better picture of the influence of the home literacy environment on reading achievement.

It is also argued that, even within the same country, the use of a home literacy instrument should be considered and used with caution. In most developing nations, the differences between urban-based families and rural or farm-based families are so great that the use of the same home literacy instrument in those different environments may not produce accurate results. For example, most urban-based children are exposed to television, have electricity, and read newspapers among other things. For rural or farm-based children, all these amenities are not even imaginable for them. So the concept of the home literacy environment between the two settings is very different. Because of this, it is also recommended that the home literacy instrument should be appropriate, not only to a country, but also to the local community in which the research study is being conducted.

REFERENCES

- Alaimo, K., Olson, C. M., and Frongillo, E. A., Jr. (2001). Food insufficiency and American school-aged children's cognitive, academic, and psychosocial development. *Pediatrics*, 108, 44-53.
- Arum, R. (1998). The effects of resources on vocational student educational outcomes: Invested dollars or diverted dreams? *Sociology of Education*, 71, 130-151.
- Astone, N.M., and McLanahan, S.S. (1991). Family structure and high school completion. *American Sociological Review*, 56, 309-320.
- Blake, J. (1989). *Family Size and Achievement*. University of California Press.
- Bradley, R. H., and Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371-399.
- Chiu, M., and Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: Country, school, and student level analysis. *American Educational Research Journal*, 42(4), 575-603.
- Christian, K., Morrison, F. J., and Bryant, F. B. (1998). Predicting kindergarten academic skills: Interactions among child care, maternal education, and family literacy environments. *Early Childhood Research Quarterly*, 13, 501-521.
- Clarke, A. and Kurtz-Costes, B. (1997). Television viewing, educational quality of the home environment, and school readiness. *The Journal of Educational Research*, 90(5), 279-285.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 94-120.

- Darling-Fisher, C. S., and Tiedje, B. (1990). The impact of maternal employment characteristics on fathers' participation in child care. *Journal of Applied Family and Child Studies*, 39, 20-26.
- Duncan, G. J., and Brooks-Gunn, J. (Eds.). (1997). *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.
- Duncan, G. J., Brooks-Gunn, J., and Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development*, 65, 296-318.
- Elder, G. H. (1974). *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press.
- Emery, R. E., and Tuer, M. (1993). Parenting and the marital relationship. In T. Luster and L. Okagaki (Eds.), *Parenting: An Ecological Perspective* (pp. 121-148). Hillsdale, NJ: Erlbaum.
- Greenstein, J. (1954). Effect of television viewing upon elementary school grades. *The Journal of Educational Research*, 48, 161-176.
- Griffin, E. A., and Morrison, F. J. (1997). The unique contribution of home literacy environment to differences in early literacy skills. *Early Child Development and Care*, 127-128, 233-243.
- Grissmer, D.W., Kirby, S.N., Berends, M., and Williamson, S. (1994). *Student Achievement and the Changing American Family*. Santa Monica, CA: RAND: Institute on Education and Training.
- Hanson, T., McLanahan, S., and Thomson, E. (1997). Economic resources, parental practices and children's well-being. In G. J. Duncan and J. Brooks-Gunn (Eds.), *Consequences of Growing Up Poor* (pp. 190-238). New York: Russell Sage Foundation.
- Heyneman, S. P., and Loxley, W. A. (1982). Influences on academic performance across high and low income countries: A re-analysis of IEA data. *Sociology of Education*, 55, 13-21.
- Korenman, S., Miller, J. E., and Sjaastad, J. E. (1995). Long-term poverty and child development in the United States: Results from the NLSY. *Children and Youth Services Review*, 17, 127-155.
- Leseman, P.P.M., and de Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly*, 33, 294-318.
- McCardle, P. and Chhabra, V. (2004). *The Voice of Evidence in Reading Research*. Bethesda, MD: National Institute of Child Health and Human Development.
- Murphy, J., Pagano, M., Nachmani, J., Sperling, P., Kane, S., and Kleinman, R. (1998). The relationship of school breakfast to psychosocial and academic functioning. *Archives of Pediatric and Adolescent Medicine*, 152, 899-907.
- National Center for Education Statistics. (2003). *The Nation's Report Card: Reading*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Neuman, S. B., and Prowda, P. (1982). Television viewing and reading achievement. *Journal of Reading*, 25, 173-183.
- Ochs, E., Taylor, C., Rudolph, D., and Smith, R. (1992). Storytelling as a theory-building activity. *Discourse Processes*, 15, 32-72.
- Ogbu, J. U. (1981). Origins of human competence: A cultural-ecological perspective. *Child Development*, 52, 413-429.
- Parcel, T. and Dufur, M. (2001). Capital at home and at school: Effects on student achievement. *Social Forces*, 79(3), 881-911.
- Parcel, T. and Menaghan, E. (1990). Maternal working conditions and child verbal ability: Studying the transmission of intergenerational inequality from mothers to young children. *Social Psychology Quarterly*, 53, 132-47.

- Payne, A. C., Whitehurst, G. J., and Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly, 9*, 427-440.
- Richie, D., Price, V., and Roberts, D. (1987). Television, reading, and reading achievement: A reappraisal. *Communication Research, 14*, 27-46.
- Saracho, O. N. (1997b). Using the home environment to support emergent literacy. *Early Child Development and Care, 201-216*.
- Scott-Jones, D. (1984). Family influences on cognitive development and school achievement: In E. Gordon (Ed.), *Review of Research in Education* (pp. 259-296). Itasca, IL: F. E. Peacock.
- Scott-Jones, D. (1987). Mother-as-teacher in the families of high- and low-achieving low-income black first-graders. *Journal of Negro Education, 56*, 21-34.
- Slaughter, D. T., and Epps, E. G. (1987). The home environment and academic achievement of black American children and youth: An overview. *Journal of Negro Education, 56*, 3-20.
- Snow, C. E., Perlmann, R., Berko Gleason, J., and Hooshyar, N. (1990). Developmental perspectives on politeness: Sources of children's knowledge. *Journal of Pragmatics, 14*, 289-305.
- Sputa, C. L., and Paulson, S. E. (1995). Birth order and family size; Influences on adolescents' achievement and related parenting behaviors. *Psychological Reports, 76*, 43-51.
- Stafford, L., and Dainton, M. (1995). Parent-child communication within the family system. In t. J. Socha and G. H. Stamp (Eds.), *Parents, Children, and Communication: Frontiers of Theory and Research* (pp. 3-21). Hillsdale, NJ: Erlbaum.
- Stevenson, D. L., and Baker, D. P. (1987). The family-school relation and the child's school performance. *Child Development, 58*, 1348-1357.
- Swick, K. J., and Broadway, F. (1997). Parental efficacy and successful parent involvement. *Journal of Instructional Psychology, 24*, 69-75.
- Third International Mathematics and Science Study Group. (1995). *Gender Differences in achievement*. Boston: Boston College.
- UNESCO-IIEP (2004). Data Archive for the SACMEQ I and SACMEQ II projects, Version 3.0.
- Walberg, H. J. (1991). Home environment and school learning: Theories, models, and evidence. Unpublished paper. University of Illinois at Chicago.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*, 461-481.
- Williams, P.A., Haertel, E.H., Haertel, G. D., and Walberg, H.J. (1982). The impact of leisure-time television on school learning: A research synthesis. *American Educational Research Journal, 19*, 19-50.

What makes a difference between two schools? Teacher job satisfaction and educational outcomes

Moosung Lee

College of Education, University of Minnesota *leex2125@umn.edu*

This article examines the interplay between school factors and teacher job satisfaction that influences educational outcomes by comparing two Non-governmental Organisation (NGO) schools in Cambodia. This small-scale qualitative case study was conducted over a period of eight weeks in 2005 in two NGO schools in a suburb of Phnom Penh. The findings show that teacher job satisfaction is crucially influenced by remunerative incentives such as salary level and welfare conditions. However, job satisfaction is also intertwined with non-remunerative incentives such as school management, principal leadership, professional development, and a meaningful sense of life through teaching. That is, both remunerative and non-remunerative incentives are associated with teacher job satisfaction. However, according to different school conditions, either remunerative or non-remunerative incentive is more prioritised by teachers. Finally, the different job satisfaction between the two faculties seems to result in the educational gap such as student enrolment rates and achievement between the schools.

Teacher job satisfaction, school factors, educational outcomes, Cambodia, NGO schools

INTRODUCTION

The purpose of this study is to investigate school factors that have a significant impact on teacher job satisfaction in non-formal primary education settings. Considering the educational situations of developing countries like Cambodia where school-related factors are more significant than non-school factors in determining student achievement, teacher quality has a crucial impact on pupils' academic performance (Mwamwenda and Mwamwenda, 1989; Lockheed and Verspoor, 1991). More importantly, teacher quality is inextricably intertwined with teachers' perception on their work life (Perry, Chapman, and Snyder, 1995). Therefore, teacher job satisfaction is often regarded as a significant determinant on the educational outcomes such as enrolment and achievement. Indeed, literature on teacher job satisfaction has consistently shown a significant relationship between teacher job satisfaction and student achievement (Heller, Rex, and Cline, 1992; Leslie, 1989). For this reason, research has continuously attempted to identify various factors influencing teacher job satisfaction (Bein, Anderson, and Maes, 1990; Clarke and Keating, 1995; Rossmiller, 1992).

Despite the previous research, relatively little attention has been paid to the discussions about teacher job satisfaction in the non-formal education settings of Cambodia, which is identified by teachers' voice. As such, it was hoped that the findings of this study would contribute to the literature on teacher job satisfaction in non-formal primary education settings in developing countries as well as advise the direction of educational policy-making.

RESEARCH BACKGROUNDS

Over the last decade, Cambodia has achieved remarkable jumps in primary school enrolments. According to the Ministry of Education Youth and Sport of Cambodia, primary school enrolments in 1996/7 were 1,918,985 and this proceeded to 2,747,080 in 2003/2004 (Bray and Bunly, 2005, p.19). To achieve this educational outcome, many NGO schools have been in active in Cambodia. As of 1999, there are more than 250 local NGOs and 200 international NGOs and most of them

are closely related to development issues including education (ADB, 1999, p.5). According to Bray and Bunly (2005), NGOs and external donor agencies provide 18 per cent of the total resources for primary education whereas the government provides 12.5 per cent of the entire costs for primary education in Cambodia. For this reason, NGO schools in developing countries like Cambodia are often regarded as crucial alternatives to overcome state failure in the provision of basic education. Indeed, recent research (Sukontamarn, 2004; Khan, 2003; Sakya, 2000) shows the positive impact on educational outcomes of NGO schools in developing countries.

Despite this general success of NGO schools in enrolment in primary education, there are various gaps in educational outcomes among them because NGO schools in Cambodia are not homogenous in terms of origin, capacity, *modus operandi* of organisation. Even the NGO schools¹ in this study that are located in the same community called Bassac Village have shown significantly different educational outcomes in enrolment rate and achievement. Furthermore, one of the schools showed frequent teacher absence during the research period.

These phenomena lead the researcher to question the school factors affecting teacher job satisfaction because the researcher noticed the outside conditions surrounding the schools are quite similar in terms of community support, community environments, and family background of pupils.²

RESEARCH METHOD

This was a small-scale qualitative case study based on a comparative analysis of the two NGO schools in a peri-urban area of Phnom Penh, Cambodia. A case study approach was used as the main research method because it was hoped that the researcher would pursue the teachers' perceptions and voices. This study also chose two NGO schools in a purposive fashion by comparable case selection in the sense that NGO schools in Cambodia are not homogenous (Goetz and LeCompte, 1984). By doing so, it was intended to show how the different school factors have an impact on forming diverse teacher views of job satisfaction.

This case study consisted of classroom observations, surveys, interviews, and focus group discussions. Based on a comparative perspective, major school factors that affect job satisfaction were compared by concrete categories.

This study involved 11 primary school teachers in the NGO schools where the researcher participated for eight weeks in 2005. No teachers were married (six females and five males). Nine out of 11 teachers finished grade 12 (one female teacher finished grade nine and one male teacher currently studies in a college). The schools are not far from the center of Phnom Penh, the capital of Cambodia. It takes at most 15 minutes to travel from the center of Phnom Penh to the schools by motorbike, which is one of the main modes of transportation of Phnom Penh. Under this situation, the researcher could visit the two schools five days a week during the first three weeks of research.

During the first week, the researcher observed classroom activities from kindergarten to grade three and conducted unstructured interviews with teachers and principals in the schools. The researcher also reviewed in-house documents such as attendance roll books, textbooks, school reports, student information report, statistical data of the schools, etc. By doing so, the researcher decided the most appropriate research methods including semi-structured interviews and focus group discussions. From the second week to the seventh week, 41 girls and 32 boys in grade three from the two schools were invited to take the same standardised test (Khmer and math)³ in order to compare student achievements between the schools. Each examination paper consisted of 20 objective questions that were made by four teachers from the schools. Based on the test results,

¹ To protect the personal information of teachers involved in this study, the two pseudonyms 'Bassac River School' and 'Naga School' were used throughout this study.

² The similarities of the community conditions surrounding the schools will be detailed in the sections of findings.

³ The students who took the test use the same textbooks, which were distributed by UNICEF.

the researcher identified that there is a big gap between the schools. Along with this, the researcher conducted two basic survey that reveals students' socio-economic status and teacher job satisfaction during the period. In addition, 22 interviews with eleven teachers (four teachers from Bassac River School and seven teachers from Naga School) were conducted. Each teacher was interviewed twice for more than an hour at a time. After the first interview with teachers, a short survey was conducted about teacher job satisfaction based on five-point Likert scale. At the same period, a focus group consisting of four teachers from Bassac River School and seven teachers from Naga School, was organised to discuss their perspectives on their job satisfaction. Focus group discussions were conducted twice for four hours and teachers' perceptions on their job were asked and discussed. During the last week of the research, the researcher obtained feedback from teachers by debriefing what the researcher found.

In interpreting Khmer, one interpreter who can speak both Khmer and Korean (the researcher's native language) was involved throughout this study. He works for an NGO, which has financially supported the Bassac village over the last three years. He also helped the researcher to understand the test papers.

FINDINGS

As stated above, out-of-school factors influencing the schools are quite similar in terms of community support, community environments, and family background of pupils. However, the schools had shown different levels of student achievement and student enrolment.

Before going further to analyse the cause of the different school performance, it is necessary to detail the similar community conditions surrounding the schools. The two schools are located in the same village. Many houses in the village are illegally built, which are not authorised by the government, as a result, many houses do not have electricity and water. Most houses were also hurriedly built because most people in the village came from rural areas to make a living or some of them are refugee from Viet Nam. Table 1 shows the similarity of the schools in terms of a family background.

Table 1. Family Background of Pupils (from Grade 1 to Grade 3)

Schools	Bassac River		Naga	
	Mean	SD	Mean	SD
Monthly income of parents including supporters like relatives and neighbors	22.05 US	6.29	20.57 US	6.13
Family Structure				
Mother and Father	47 (72.3%)		116 (78.3%)	
Single father	1 (1.5%)		2 (1.3%)	
Single mother	8 (12.3%)		21 (14.2%)	
Orphan supported by neighbors or relatives	9 (13.8%)		9 (6.1%)	

Source: The data about family structure were retrieved from in-house documents (Bassac N = 65 and Naga N = 148). A survey was conducted to obtain the data about the income of parents and 21 pupils' survey was not retrieved (Bassac N = 52 and Naga N = 140). US \$ is also used as a main currency in Cambodia (1 US \$ = 4,000 Riel).

Based on a t-test, there was no significant difference in the monthly average income of parents between the schools. In terms of family structure, most of the pupils live with their parents, but approximately 28 per cent of Bassac pupils and 22 per cent of Naga pupils suffer the loss of their father or mother. According to the principals of the two schools, most of pupils' parents are temporary daily workers (for example, construction workers), motorbike drivers, self-employed (for example, grocery store, water store, and stationery store), and vendors (for example, books, food, shells, and clothes). Also, some of them are unemployed and only a few are regular workers. Although there was no official data about the length of education of the parents, both teachers in the schools pointed out that most of pupils' parents are illiterates. This implies that the length of education of the parents is likely to be less than six years. In addition, as far as community environments are concerned, there was no crucial difference between the schools because the Bassac village is a poverty-stricken village where most people live in miserable conditions on less than one dollar a day. Not surprisingly, most parents of the pupils enrolled in

the schools can not afford to send their children to public schools. At most 10 per cent of the school-age children in the village are enrolled in the public schools.⁴ Thus, it can be argued that the outside conditions surrounding the two schools are almost the same.

Under this situation, most people take the lack of resources to support the schools for granted at the community level. Even though Bassac River School was established by the governing body of the community, there was no particular fund allocated for the school because the governing body heavily relied on external donors. In the case of Naga School, since it was set up by an external religious organisation, there was no community connection as well. Rather, Naga School pays \$75 USD for rental fee per month to the governing body of the community for the land of school buildings. In this context, the schools are expected to be financially self-sustaining. For these reasons, both the schools have internal regulations in managing the schools, which are quite free from the regulation of the government's educational policy. In brief, all these conditions allowed the researcher to concentrate on school factors.

Interestingly, despite the similar community conditions surrounding the schools, the two schools showed strikingly different school performance in terms of gender equity. Based on the investigation of student attendance books, significant differences of girls' drop-out rate and enrolment rate were identified.

Table 2. Enrolment and Drop-out Rates (from Grade 1 to Grade 3)⁵

	Bassac River School		Naga School	
	Girls	Boys	Girls	Boys
School enrolments	25	40	80	68
School enrolments rate of girls to boy	62.5%	100%	117.6%	100%
Drop-out pupils (up to Grade 3)	18	26	14	10
Rate of drop-out pupils to school enrolments	72.0%	65.0%	17.5%	14.7%

As seen in Table 2, there was an obvious gap in both school enrolment rates of girls and the number of drop-out between the two schools. Compared to the school enrolment rates of girls to boys in Bassac River School (62.5%), Naga School shows a higher school enrolments rate of girls to boys (117.6%). Additionally, girls who could not complete their third grade in Bassac River School were surprisingly 72.0 per cent of the total girls, compared to 17.5 per cent of Naga School. In brief, Naga School showed a better school performance than Bassac School in terms of both enrolment and drop-out rates of girls.

Furthermore, there was a significant achievement gap between the two schools. 72 pupils (21 pupils from Bassac River School and 51 pupils from Naga School) in grade three from the two schools were invited to take the same standardised examinations (Khmer and math) to compare pupils' achievements.⁶ As illustrated in Table 3, there was a significant difference of student achievement between the schools. The result shows that the pupils in Bassac River School lag behind the peers in Naga Schools in terms of both Khmer ($t = -4.458, p < 0.001$) and math ($t = -4.010, p < 0.001$).

Table 3. Student Achievement between the Schools

	Bassac River School (N=21)			Naga School (N=51)			t	Sig. (2-tailed)
	M	SD	SE	M	SD	SE		
Khmer	55.95	9.70	2.11	66.57	8.97	1.25	-4.458	.000
Math	48.57	10.97	2.39	58.24	8.54	1.19	-4.010	.000

There was another gap related to teachers. During the second to the third week of the research, a short survey about job satisfaction was conducted. It was based on a five-point Likert scale (1

⁴ This estimated measure is based on the interview with the principals of the two schools.

⁵ Both of the schools provide a curriculum from kindergarten to grade three. In this table, the 50 kindergarten pupils in the two schools were excluded because this study focused on primary education. Also, the drop-out pupils here means pupils who did not attend the schools more than 6 months once they enrolled the schools.

⁶ Originally, 43 girls and 32 boys were enrolled in the third grade in the schools but three pupils were absent from the schools when the pupils were invited to take the test.

never satisfied and 5 very much satisfied). The result represented that the job satisfaction of Naga School teachers was much higher than that of Bassac River School teachers: Bassac River School (M 1.75 and SD 0.95) and Naga School (M 4.57 and SD 0.53).

DISCUSSION: WHAT MAKES A DIFFERENCE?

Then, what school factors make a difference between the schools? Table 4 shows the similarities and differences of school factors between the schools. In particular, the shaded parts indicate the differences between the schools that possibly make a difference in forming teacher job satisfaction.

Table 4. School Factors (Physical Conditions and Management)

Factors	Sub-factors	Bassac River School	Naga School
Physical Conditions	School size (by enrolment)	100 (Girl 40, Boy 60)	173 (Girl 89, Boy 84)
	Teacher-Student Ratio	1: 25.0	1:24.7
	Age of enrolled pupils	Average 8.26 Range 4 -14	Average 9.39 Range 4 -14
	Classroom	2 classrooms, 1 multi-purpose room	3 classrooms, 1 multi-purpose room
	Other facilities	1 teachers' room shared with the administration office of Bassac Village)	1 small front yard for school assembly
	Major classroom facilities	Desk, chair, teacher cabinet, and whiteboard	Desk, chair, teacher cabinet, whiteboard, and fan
	Quality of facilities	Very worn out building made by wood	Shabby tin-roofed buildings
Management	School origin	Established by community and external donors	Set up by a external religious organisation
	Grade	From kindergarten to Grade 3	From kindergarten to Grade 3
	PTA	Nominal PTA	None
	School expenditure per month	Data not accessed	\$960 USD (on average)
	Tuition fee and materials fee	Free	Free
	Free meals	No school meals	School meals provided twice a day
	Independence from central government	Independent	Independent
	Admission criteria	Admit by residence	Admit by residence
	Number of school shifts	1	2
	Number of class days	5 days a week and 2 vacations (about 90 days)	6 days a week and no vacation and extracurricular activity for street children every Friday late afternoon
	Educational background of principal	Primary school graduate	University graduate
Degree of principal emphasis on pupils' attendance	Strong but the first priority in school management is to draw fund from external donors	Very strong and the first priority in school management	
Principal leadership	Authoritarian leadership	Authoritative leadership	

In terms of the physical conditions of the schools, although there are some differences in the number of classrooms, particular facilities (teachers' room and front yard), and enrolments between the schools, those conditions did not make a difference in shaping job satisfaction.

Considering the circumstances of Bassac Village, the facilities of school are not bad. (Teacher 1, Naga School)

That is, the worn-out physical conditions of their school are evident problems but it did not have a significant impact on teachers' motivation. Most teachers in both schools complained about the lack of physical resources, which was continuously highlighted by teachers during the research, but also they were already accustomed to teaching in this situation.

My school always lack materials like stationery and facilities like toilets. Obviously those things are inconvenient for teaching and learning. But also we have been always accustomed to this situation (Teacher 3, Bassac River School).

However, the conditions of school management made a difference teacher job satisfaction between the schools. Ostensibly, Bassac River School seemed to take advantage of community support because it has a PTA.

Nonetheless, the school has a difficulty in promoting the residents of Bassac Village to join its activities because most residents are struggling for a living. Under this lack of interest and involvement from the community, the PTA was an organisation in name only. This situation also influenced the principal leadership of Bassac School. That is, all decision power was given to the principal. Not surprisingly, the PTA does not have any legislative or advisory activities. For this reason, some of the teachers in Bassac River School complained about the management by the principal.

The principal emphasises collateral affairs, for example, reception for donors/visitors or some paper works, more than the everyday life of teachers and our pupils (Teacher 1, Bassac River School).

The apathy of the principal also made the Bassac teachers feel frequently frustrated in managing school affairs. The teachers believed that their voices and perspectives of education have been neglected.

I feel like my opinions about administration and management are always ignored (Teacher 3, Bassac River School).

Thus, the teachers of Bassac viewed that their principal should pay attention to education itself and should be learned about “what good education is” (Teacher 3, Bassac River School). This connotes that there is a gap about the perception of good education between the teachers and the principal. Furthermore, the teachers felt that they are usually treated as passive implementers according to the principal’s decision.

In the case of Naga School, the teachers work six days a week without a vacation throughout year. In addition, Naga School runs two shifts to accommodate 173 enrolments. In these respects, it seemed that Naga teachers work harder than Bassac teachers. However, six out of seven Naga teachers were satisfied with their current working conditions because they think that they have autonomy to some extent in conducting their duty. Like the Bassac principal, the principal of Naga School also made decisions all about school management such as number of class days a year, teachers’ salary level, school shifts, school meal programs, teacher appointment, etc. However, the principal never supervised teaching styles or teaching methods of the teachers. As such, developing curriculum was thought of as teachers’ role in Naga School. These seemed to make the teachers feel relative autonomy in their job.

Most important decisions have been made by him alone, but it is his own part. My responsibility is to teach and care well for my pupils (Teacher 4, Naga School).

In addition, the Naga principal is financially supportive for the teachers’ day-to- day life and put pupils’ attendance first in school management. More importantly, transparent school spending clarified by the Naga principal seems to encourage the teachers to work hard. In fact, Naga School spends on average 960 US dollars per month in school management.⁷ Hence, the style of principal leadership should be noted as an important factor for understanding teacher job satisfaction. In brief, their job satisfaction is crucially intertwined with school management. This

⁷ According to the Naga principal, on average \$960 USD are spent a month in school management: teachers’ salary (\$310 USD), teachers’ accommodation and food (\$255 USD), free school meals (\$200 USD), utility fees for School (\$20 USD), rental fee for the land of three classrooms (\$75 USD), traffic expenses for teachers (\$70 USD), and scholarship for teachers (\$30 USD).

means that non-remunerative factors such as management style, principal leadership, and administrative procedure are also important to teacher job satisfaction.

Table 5 presents other school factors related to job satisfaction: classroom activities and welfare conditions. In terms of classroom activities, there was no significant difference between the schools. The only difference was whether there is a co-operative activity among teachers such as team teaching. The teachers in Naga School did co-operative work for the extra-curricular activities for their pupils and street children who are either orphans or poor children. Every Friday late afternoon, they taught Khmer, songs, dances, painting, and the Bible. Also, they provided food and basic medical service on a rental ship. To this end, they had a regular meeting at the teachers' dormitory and developed a special curriculum and learning materials. This extra-curricular activity played an important role in gathering other teachers' experience and knowledge about teaching. Through this co-operative work, they "understand their colleagues" and "feel they are contributing to something meaningful values" through their job (Teacher 6, Naga School).

Table 5. School Factors (Classroom Activities and Welfare Conditions)

Factors	Sub-factors	Bassac River School	Naga School
Classroom activities	Instructional materials	Textbooks (distributed by UNICEF), exercise book, and chalkboard	Textbook (distributed by UNICEF), exercise book, and chalkboard
	Instructional time	50-minute class time and 10-minute break (08:00 – 10:00)	50-minute class time and 10-minute break (08:00 – 11:00 and 14:00 – 17:00)
	Major pedagogy type	Whole class teaching	Whole class teaching
	Co-operative work for classroom activities	None	Once a week for extra-curricular activities
	Frequency of testing pupils' achievements	Once a month	Once a month
Welfare conditions for teachers	Salary level	No salary, only basic material support (e.g., food, accommodation, and transport allowance)	Salary from 35 to 55 dollars based on years of teaching besides 25 dollars per teacher for allowance
	Transport for teacher	1 motorbike shared with the village office	2 motorbikes only for teachers
	Other welfare systems	None	Dormitory and scholarships for teachers
	Opportunities of in-service training	None	None
	Job security and status	Volunteering, temporal, and part-time job	Full time job

Another distinctive school factor affecting teacher job satisfaction was welfare conditions for teachers. Obviously, salary was the most important incentive for teachers' morale in the schools. Related to this, the teachers of Bassac River School complained about their job status.

I am basically volunteering as a teacher. Although the village office provides basic materials such as transport allowance, food, and temporary accommodation located in the village. But those things are not enough (Teacher 3, Bassac River School).

The principal regards me as only a volunteering teacher, not a real teacher (Teacher 1, Bassac River School).

The difference between "a volunteering teacher" and "a real teacher" is not only semantic but also substantive in Bassac teachers' work life. Then, why do Bassac teachers continue to work as a volunteer teacher? Despite no official salary for the teachers, they had their own reasons to maintain their job status. In the case of male teachers, they believed that teaching experience in Bassac River Schools will be helpful to transfer to other jobs in the government, NGOs, or a paid job at other non-formal school in the future. Unlike their male colleagues, the female teachers in Bassac River School viewed their job as somewhat a symbolic status rather than a practical one.

In other words, though they did not get paid, they accept their position as a “decent profession” because the labor market for females in Cambodia is quite underdeveloped. Noticeably, the Bassac teachers overlooked the significant educational issues such as in-service training, curriculum, pupils’ learning, or teaching methods because they were dominated by the issue of salary.

In contrast to the Bassac teachers, the Naga teachers were satisfied with their salary level and their job security.

Compared to other jobs such as police or public school teachers, my salary is not small. Also, public school teachers have to sell learning materials to pupils and force pupils to have private tutorials in order for teachers to make a side income. But I don’t need to do that (Teacher 1, Naga School).

Because of this different job status, there was a clear difference about future career between the teachers. Three out of four teachers of Bassac River School regard their current job as a temporal one. They did not consider teaching itself as their lifelong career. In contrast, five out of seven teachers at Naga School desired to teach continuously and the teachers also hoped to get a degree or a license related to teaching profession.

Within this context, when the teachers in the two schools were asked “what do you need first for teaching better?,” the teachers answered differently as well. Three out of four Bassac teachers stated that they need more material resources and support including teachers’ salary from the village. In the case of Naga School, five out of seven teachers require themselves to be professionally trained as a teacher by obtaining opportunities for in-service trainings and degree programs. The five teachers believe that there is no professional teacher at Naga School and thus they need more knowledge about teaching methods to “be good teacher.” In brief, whereas the Bassac teachers see salary as the most important satisfaction factor, the Naga teachers view opportunities for career development as the most crucial satisfaction factors along with salary.

CONCLUSIONS

The teachers in both of the schools regard salary as the key remunerative factor affecting their job satisfaction. School management is also a more important factor of teacher job satisfaction than physical facilities of schools. Despite these similar factors affecting teacher job satisfaction, there were two distinctive features regarding teacher job satisfaction. One is that Bassac teachers seemed to be based on more instrumental satisfaction factors whereas Naga teacher seemed to be oriented in self-growth factors. In other words, Bassac teachers placed more emphasis on remunerative incentives. They did not mention about profession development as an element of their job satisfaction because most of them did not view teaching profession as their lifelong career. In contrast, Naga teachers mentioned more multiple satisfaction factors than Bassac teachers: in-service training, scholarships for further degrees, opportunities for teaching certificate, job status, and a sense of meaningful life as a teacher. In conclusion, it is evident that teacher job satisfaction is basically influenced by remunerative incentives such as salary level and welfare conditions. However, it is also closely intertwined with non-remunerative incentives such as school management, principal leadership, professional development, and a sense of meaningful life through teaching. That is, both remunerative and non-remunerative incentives interplay in determining teacher job satisfaction. Importantly, based on their different school conditions, either remunerative or non-remunerative incentive was more highlighted by teachers. Intuitively, this different job satisfaction between the two faculties seems to be closely related to the educational gap between the schools. These findings are sufficient as a first approximation for the understanding of teacher job satisfaction in the context of Cambodian NGO schools. Finally, as teacher job satisfaction seems to form different teacher professionalism in the two schools, through a follow-up study the researcher will further analyse how teacher job satisfaction is associated with forming diverse discourses of teacher professionalism.

REFERENCES

- Asian Development Bank. (1999). *A study of NGOs: Cambodia*. [Online] <http://www.adb.org/NGOs/docs/NGOCambodia.pdf> [2005, August 15].
- Bray, M. and Bunly, S. (2005). *Balancing the Books: Household Financing of Basic Education in Cambodia*. Hong Kong: Comparative Education Research Centre, University of Hong Kong.
- Bein, J., Anderson, D. and Maes, W. R. (1990). Teacher Locus of Control and Job Satisfaction. *Educational Research Quarterly*, 14 (3), 7-10.
- Clarke, R. and Keating, W. (1995). *A Fresh Look at Teacher Job Satisfaction*. [Online] http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/24/fe/37.pdf [2005, August 10]
- Goetz, J. and LeCompte, M. (1984). *Ethnography and Qualitative Design in Educational Research*. New York: Academic Press.
- Heller, H. W., Rex, J. C., and Cline, M. P. (1992). Factors Related to Teacher Job Satisfaction and Dissatisfaction. *ERS Spectrum*, 10 (1), 20-24.
- Khan, S. (2003). Participation via Collective Action in Government and NGO Schools in Pakistan. *Development in Practice*, 13 (4), 361-376.
- Leslie, K. (1989). Administrators Must Consider and Improve Teacher Satisfaction. *NASSP Bulletin*, 73, 19-22.
- Lockheed, M. and Verspoor, A. (1991). *Improving Primary Education in Developing Countries*. Oxford: Oxford University Press.
- Mwamwenda, T., and Mwamwenda, B. (1989). Teacher Characteristics and Pupils' Academic Achievement in Botswana Primary Education, *International Journal of Educational Development*, 9 (1). 31-42.
- Perry, P., Chapman, D., and Snyder, C. (1995). Quality of teacher worklife and classroom practices in Botswana, *International Journal of Educational Development*, 15 (2). 115-125.
- Rossmiller, R. A. (1992). The Secondary School Principal and Teachers' Quality of Work Life. *Educational Management and Administration*, 20 (3), 132-146.
- Sakya, T. (2000). Role of NGOs in the Development of Non Formal Education in Nepal. *Journal of International Cooperation in Education*, 3 (2). 1-17.
- Sukontanarn, P. (2004). *The Entry of NGO Schools and Girls' Educational Outcomes in Bangladesh*. [Online] <http://econ.lse.ac.uk/phdc/otherfiles/sukontamarn JM paper.pdf> [2004, December 3].
- Yin, R. K. (1989). *Case study research: Design and methods*. Thousand Oaks: Sage.

Why do they not talk? Towards an understanding of students' cross-cultural encounters from an individualism/collectivism perspective

Joanna K.L. Tan

Student Services Centre, Nanyang Technological University, Singapore joannatan@ntu.edu.sg

Jonathan W.P. Goh

National Institute of Education, Nanyang Technological University, Singapore
jonathan.goh@nie.edu.sg

Many universities today are promoting cultural diversity to prepare students to be competent in intercultural communication and function effectively in an increasingly global world. The purpose of this study is to explore how and why students from different cultural backgrounds are motivated to communicate and interact with each other. It is based on the premise that a given culture prescribes how and why individuals are motivated to interact with others from different cultural backgrounds. In order to provide a holistic view that would account for the multi-faceted and subtle nuances of human behaviour in communication, a naturalistic inquiry approach was adopted to explore the communication patterns of groups of students from predominantly individualistic and collectivistic cultures. The findings indicate layers of differences in the way students relate across cultures in class as well as social environments. The implications for inter-cultural communication for educators are discussed using the individualism-collectivism dimension.

Cross-cultural encounters, communication competence, individualism and collectivism, cultural values, ethnography

INTRODUCTION

As the world becomes increasingly interconnected and culturally diverse, the internationalisation of education has become a major goal of many universities. Through the promotion of cultural diversity, universities endeavour to prepare individuals to become competent in intercultural communication and function effectively in a multicultural environment (Albert and Triandis 1985; Blight 1997; Nowak, Weiland and McKenna 1998). Indeed, the growth of cultural diversity on university campuses has resulted in a plethora of interest and research in the field of intercultural communication internationally. Interestingly, however, a review of the literature suggests that although universities continue to celebrate the cultural diversity of their student population, cross-cultural interaction among students remains alarmingly low (for example, Burke 1986; Schram and Lauver 1988; Volet and Pears 1994, 1995; Mullins, Quintrell and Hancock 1995; Nesdale, Simkin, Sang, Burke and Fraser 1995; Todd and Nesdale 1997; Volet and Ang 1998). As such, Bochner, Hutnik and Furnham (1986, p.5) assert that it would be naïve to assume that “close proximity would spontaneously result in individuals from different cultures interacting and forming long-lasting friendships”. Instead, several studies on the formation of academic or social networks, have found that students are constantly forming support groups, which are heavily ethnic-based, with very little mixing of cultures (for example, Tanaka, Takai, Kohyama, Fujihara and Minami 1997; Nowak, Weiland and McKenna 1998).

The aim of this study is to explore how and why individuals are motivated to communicate inter-culturally. It is based on the premise that culture and the motivation to communicate are fundamentally and ineluctably linked. Samovar and Porter (1991: 47) argue that “people learn to think, feel, believe and strive for what their culture considers proper.” Indeed, the dominant cultural norms and rules have an effect on the attitudes and behaviours of individuals toward what is desirable (Albert 1970; D’Andrade 1992; Hall 1977; Schneider 1976; Vogt and Geertz 1973).

A communication encounter (i.e. both verbal and non-verbal) is understood here as being a context where two individuals meet and bring with them different perceptual positions regarding levels of communication and life experiences, which are in turn influenced by the individual’s cultural background (Westwood and Borgen 1988). To achieve intercultural communication competence, it would require the accurate transmission of the expectations and experiences of each individual to result in mutual understanding. In short, interaction not only involves the perception, selection, and interpretation of salient features of the rules used in actual communicative situations, but also the integration of these with the communicator’s motivation, cultural knowledge and skills (Porter and Samovar 1991).

HOW I INTERACT WITH OTHERS? THE INDIVIDUALISM VERSUS COLLECTIVISM DIVIDE

The individualism-collectivism construct has been widely researched as a cultural construct to investigate the nature of individuals, society as well as the relationship between them (for example, Rokeach 1968; Hall, 1977; Hofstede, 1980; Bond, Leung and Wan, 1982; Triandis 1995; Schwartz 1994; Schwartz and Sagiv 1995; Hofstede and McCrae 2004). Essentially, individualism and collectivism are two views of the self which affect how an individual relates to others. Broadly, individualists view themselves as agents of their existence and primary focus is placed on the needs of the individual. This view does not deny the existence of the society. Rather it holds that society is made up of a group of independent individuals, not something over and above them. Conversely, collectivists focus on the needs of the group and identify themselves as an extension of other individuals in that group (for example, family, community, society or nation). Rather than deny its existence, the individual’s identity is defined by its relationship with members of the group.

In their study, Markus and Kitayama (1991) were particularly interested in exploring how such different views of the self affected communication styles and the resulting consequences. Consistent with other researchers, Markus and Kitayama (1991) argue that individualism can be found in many Western cultures. This view of the self focuses on the internal attributes of the individual and views the individual as an autonomous, independent person. Thus, in interpersonal relationships, the independent self is more likely to seek to assert his or her inner characteristics (such as personality, desires, personal goals and emotions) in order to change the social situation. Other labels used to describe the independent or individualist self include, low-context, egocentric, separate, autonomous, idiocentric, self-contained, openness to change, and self-enhancement.

In contrast, many non-Western cultures, such as the Chinese and Japanese, tend to view the self as an extension of significant others including family members, friends and co-workers. This view of the self maintains that individuals are interconnected, and depend on each other for self-definition (Shweder and Bourne 1984; Chu 1985; Hsu 1985, King and Bond 1985; Gao 1998). Thus, an interdependent individual sees the self as less differentiated from others and is more likely to find ways to fit in with others. In other words, interdependent individuals would seek to form obligations, and become part of various interpersonal relationships. As such, the interdependent self is also more likely to attempt to alter or control his or her inner attributes (i.e. behave in a modest, self-effacing, or other-enhancing way) to suit the context and avoid disrupting harmony in the relationship (Yang 1992). This is due to the belief that in abasing

oneself or actively enhancing the position of others, one would in effect enhance the self (Markus and Kitayama 1991). Modest behaviour is not viewed as a sign of weakness or a failure to express oneself; rather, it reflects tolerance, self-control, flexibility and maturity. The interdependent or collectivist view of the self is variously referred to as sociocentric, holistic, allocentric, ensembled, constitutive, contextualist, high-context, connected and relational, conservatist, self-transcendent.

Generally, South-East Asian societies like Singapore and Malaysia are associated with collectivist orientation while Western societies, like Australia, are associated with individualism (Triandis 1995). In reality the two extreme cultures do not exist in their pure form as represented. However, such a dichotomy allows researchers to categorise and explain cross-cultural similarities and differences in a more explicit and easily comprehensible way (Kluckhohn and Strodtbeck 1961; Das 1972; Triandis 1995). Kluckhohn and Strodtbeck (1961) used the concept of whether the relationship between individuals and groups places the individual or group as the focus to formulate their *linear* (for example, submission to elders), **collateral** (for example, agreement with group norms), and **individualistic** (for example, doing what self-conceptions dictate) orientations. Similarly, it is Schwartz (1994)'s cultural dimensions of **Self-Enhancement/Openness-to-Change** (in other words, Individualism) versus **Self-Transcendence/Conservation** (that is, Collectivism) form a continuum of related motivations rather than distinct ones. This continuum gives rise to a circular structure where the definition of value types represents conceptually driven decisions about where one hazy set ends and another begins (Schwartz and Sagiv 1995, p.94). Therefore, it is not surprising if individuals in all cultures hold both individualistic and collectivistic views at the same time, albeit of varying strengths (Ho 1988).

METHOD OF INQUIRY

A naturalistic inquiry approach was employed in order to provide a holistic view that would account for the many facets of human behaviour in communication (Munro 1997). In particular, significant focus was placed on the interactants' perspectives in order to reveal the particularistic and subtle nuances of interpersonal relationships across cultures. This revelation is important because part of the process of attaining competence in intercultural communication is in being able to understand and appreciate the different indigenous culture contexts and to take that into account in interacting with others.

The chosen field site was a university in Australia with an international student population on campus of about 1200, or 8.5 per cent of the student population. Of this group, 42 per cent were from Singapore, 28 per cent from Malaysia, and 18 per cent from other parts of Asia. The majority of these students were of ethnic Chinese background. In total, over a year was spent in conducting fieldwork at the university.

During fieldwork, both participant and non-participant observations were employed. In addition, the researchers spent between 2 to 4 hours interviewing each of the respondents. These semi-structured and informal sessions were used as a way to affirm the expressive power of language and allow the researchers to appreciate the cultural connotations and denotative meaning of what the respondents say (Fetterman 1989). Students were asked to relate specific intercultural communication experiences that were particularly memorable (that is, both satisfying and dissatisfying). Asking students to recount experiences that were of significance was essential it allowed them an opportunity to talk at length, volunteer information, and pursue and elaborate on issues that were of importance to them (Eggerton and Langness 1974). In addition, these questions helped refine and expand knowledge of the cultural groups and their perceptions of communication competence.

Sampling Plan

In total, 104 students were interviewed in this study. Of this group 49 were Singaporean, 15 Malaysian and 40 Australians. The characteristics of this sample that were taken into consideration were similar ethnic identification, age group, the course they were enrolled in and their level of intercultural communication experience. Most of the students were aged between 17 and 25 years old. These students were evenly distributed across the Arts and Social Sciences, Commerce, Engineering and Science faculties. For the international student group, the period of time they spent in Australia was also taken into consideration. This was because studies have shown that student's cultural orientation could change within the duration of less than six months in a foreign country (Volet and Renshaw 1993; Volet, Renshaw and Tietzel 1994).

Cultural Groups and Labels for the Groups in this Study

In order to generate valid and appropriate descriptions, the researchers employed ethnic labels that the respondents understood and preferred (Mirande and Tanno 1993). In order to achieve both contextual and representative validity, researchers focused on labels and identities commonly acknowledged by group members, as well as those ascribed to them by outsiders, specifically in intercultural communication contexts (Collier and Thomas 1988; Hecht, Collier and Ribeau 1993). For the international student groups, students from Malaysia and Singapore were considered. These students referred to themselves variously as Malaysians, Singaporeans, Asians and Chinese. It should be noted that although the majority of them acknowledge similar ethnic backgrounds that can be traced back to China, some Singaporean and Malaysian Chinese students consider themselves as quite distinct from each other. They indicate differences in outlook on life (including work and leisure) perceived rate of modernisation, as well as language differences, as distinguishing factors. Therefore, while it would be valid to assume that the Chinese in general hold largely similar cultural values, that is Confucian heritage values (see Bond 1996), it would be interesting to explore the ways in which they might differ. With reference to the Australian group, the majority of the students considered were of Anglo-Celtic background. Many of the students (both international and local Australian), however, acknowledged the fact that the Australian student population was not homogeneous and that many Australians were from Asian or non-Anglo-Celtic background. The following quotes from two Australian respondents illustrate the point,

The Asian population who are here as international students and those who have been born here and grew up here and who are Australians, there is a huge difference, these two populations. (Australian female, No. 11)

Like my friend Tong, he's Asian and he's brought up here. Like a lot of Asian students take more effort, care in their presentation, cars and clothes, brands. But you have Tong on the other hand, okay, drinking, casual clothes, pony tail, laid back, relaxed. (Australian male, No. 03)

Those who were not of Anglo-Celtic background (one student had a Singaporean Chinese mother and Irish father, another had Croatian parents, and the parents of another student were from Zimbabwe of Indian background), nevertheless, identified themselves more with the Australian Anglo-Celtic students. They attributed this to being second- or third- generation Australians and having gone through the Australian educational process. Below is a quote from the girl with Chinese and Irish parents,

Well, I grew up here and everything, attended school here. I see myself as definitely Australian. Neither of them (mum or dad) have really pushed anything on me. They

brought me up the stereotypic Aussie way. Definitely not Chinese and not Irish. (Australian female, No. 38)

FINDINGS

The results presented in Table 1 indicate that differences exist in the ways the groups of students view competent communication. For example, the majority of students from all three groups agreed that **Sense of Affiliation** was a criterion for successful encounters. However, to the Chinese students, this sense of affiliation was very much motivated by a need for security resulting from group identification. In contrast, the Australian students tended to place great value on their individuality.

Table 1. Key characteristics for intercultural communication competence

Values	Characteristics	Cultural Groups					
		Australian (N = 40)		Singaporean (N=49)		Malaysian (N=15)	
		No.*	% **	No.*	% **	No.*	% **
Individualism/ Independent Self	Sense of Affiliation	31	77 %	43	88 %	14	90 %
	Abasement	0	0	24	49	10	67
	Cultural Similarity	26	65	31	63	14	90
	Handling Cultural Taboos	4	10	4	8	5	33
	Friendliness/Sociability	16	40	15	31	14	90
	Empathy	16	40	22	45	8	53
	Self-Effacement	0	0	30	61	14	90
	Honesty	26	65	0	0	0	0
	Individuality	38	95	1	2	0	0
	Initiative	20	50	24	49	12	80
Collectivism/ Interdependent Self	Spoken Language	31	78	38	78	15	100
	Loyalty	4	10	0	0	0	0
	Negotiation Styles	0	0	24	49	12	80
	Non-Judgmental	4	10	0	0	0	0
	Having Own Opinions	26	65	4	8	2	13
	Reciprocity	3	8	24	49	14	90
	Reliability	11	28	24	49	9	60
	Role Responsibility	0	0	47	96	15	100
	Security	7	18	43	88	15	100
	Sincerity	0	0	36	74	11	73
Trust	0	0	24	49	11	73	
Upbringing	18	45	24	49	12	80	

NB: * Codings of number of responses

** Percentage of respondents

From the ethnographic data, it is evident that the students' sense of self affected their communication styles. The Chinese students tended to view the self as closely associated with and affected by other individuals who were considered significant to them. As a result, their communication styles involved strong consideration for the feelings and reactions of others, in particular the maintenance of self-face, other-face and mutual-face. The Australian students also acknowledged the value of group identification. In fact, 77 per cent of Australians, 90 per cent of Malaysians and 88 per cent of Singaporeans agreed that a **Sense of Affiliation** was important for successful communication. However, the Australian students were more likely to place importance on having a sense of **Individuality**, which dominated their communication styles (95% of Australians; 0% of Singaporeans; 2% of Malaysians).

The Chopstick Culture: "Sticking with people of the same kind"

The ethnographic data revealed that both the Australian and Chinese students had a tendency to mix with peers from similar cultural backgrounds, whether it was at work or in their social life, because of a sense of affiliation to each other. However, while the Australian students tended to

focus more on common interests, the Chinese students cited feeling a sense of *security* as an important reason for them to stick with peers from the same cultural (Chinese or Asian) background. All the Malaysians and 88 per cent of Singaporeans brought up *Security* as a key communication characteristic as compared to only 18 per cent of Australians,

I mix around with more Asians because you feel some kind of security. To me I feel comfortable because like home, like a group of people I know of [sic]. (Malaysian male, No. 10)

If that person is an Asian, that already makes you more comfortable talking. Roughly you know what to say. It's just how you feel. You can't pinpoint... just feel comfortable. (Singaporean male, No. 07)

Some of the Chinese students acknowledged that this strong group bonding among the Asian student community, which fostered the notion of an in-group and out-group, inhibited communication with the Australian students,

We are very groupish and it is very difficult to breakaway from established norms. If someone from our group tries too hard to associate with the Aussies, he would be ostracised. It is like a reverse form of racism. (Singaporean female, No. 49)

“What talk you? - You don't speak my language”

From Table 1, 78 per cent of Australians, 100 per cent of Malaysians and 78 per cent of Singaporeans identified the ability to speak a common **Language** as one of the reasons for successful communication. Many of the informants expected that speaking a common language would allow individuals to relate to and identify with each other,

I guess language is one of the main barriers that prevent us [the Australian and Asian students] from communicating with one another. (Australian male, No. 14)

I know of Australian friends who are successful in breaking into our Asian friendship circles. Like this friend of mine can speak Indonesian and other Asian language as well. It really helps. (Singaporean female, No. 12)

Clearly, speaking a language requires more than just grasping its grammar. It involves a proper way of speaking, including the accent, slang and the jargon used. This applied especially to the Chinese students from Singapore and Malaysia. Many of these students were able to speak in languages other than English (such as, Mandarin, other Chinese dialects and Malay), and often combined these languages to form a unique Singaporean Creole (that is, *Singlish*),

Although they [the Australians] speak English, their English is like more Australian slang. So I find it difficult to catch. (Malaysian male, No. 15)

In addition, even though most of the Chinese students spoke English fluently, many of them often slipped into their own languages. The reason given by some was that they did not feel as competent in English as the Australian students. Others suggested that speaking in the Australian way (or not speaking *Singlish*) was unnatural, and would set the person apart from other Asian students,

Specially when speaking up in class they [the Australian students] have the upper hand because they are better equipped than us in English and grammar. (Malaysian female, No. 04)

We can't stand Chinese that come here and they adopt the Australian slang, 'buay tah han' [Cannot stand]. ABCs [Australian Born Chinese], cannot blame them, they grew up here. But not those who just come here to study. (Singaporean male, No. 03)

In particular, speaking in *Singlish* was seen as a form of cultural identification and affiliation for many of the Chinese students and communication often took on a more informal mood when they did. However, many of the Australian informants felt that speaking in a foreign language (or a language that they did not understand) was an act of exclusion. They commented that people who chose to go to a foreign country should at least attempt to speak the local language, albeit poorly. As such, since the Asian students made a choice to come to an English-speaking country such as Australia, they should take the initiative to speak in English,

I feel that the Asian students make a big mistake by talking in their own language. It immediately sets them apart. (Australian male, No. 28)

I do find it quite offensive sometimes because we are in as such an English-speaking country and they should try to include everyone. Like if someone came into my group of friends, I'd try to include them, you know. (Australian female, No. 30)

***Guanxi* – “A Network of Friends”**

The data revealed that the Australian and Chinese students had many similarities in their views on friends. Both groups agreed that friends were people whom they could trust and rely on in times of trouble. Friends were also people with whom they shared activities, such as going to the movies, shopping, parties, restaurants for meals, pubs (mainly for the Australian students), and talking and confiding in each other. However, there were some subtle yet significant differences in how these students conceptualised friendship. For instance, many of the Chinese students tended to view friendship as a form of ‘transaction.’ They acknowledged the possibility that the relationship would allow the receiving and returning of obligations before they would enter it. In fact, 90 per cent of Malaysians and 49 per cent of Singaporeans compared to only 8 per cent of Australians brought up **Reciprocity** as a key characteristic for successful communication. Thus, a relationship that did not allow for interdependence and mutual benefit would not be highly pursued,

Somehow, I get the feeling that they [the Australians] are not willing to get to know us beyond a superficial level. It's like beyond that, they are not comfortable. They are like if you go drinking with them, there's just superficial talk. Becoming friends and all, somehow I feel it's something resisting on their part. With friends, you call each other up regularly and you really talk serious stuff and you are willing to open up. (Malaysian female, No. 12)

Getting to know Australians is not on my agenda. Only if I have to work here then I'll be interested, because if I have to work here, I have to know and understand Australians. (Singaporean male, No. 41)

However, once friendship was established among the Chinese students, it was observed that it would be highly personal, and both parties would continue anticipating the receiving and returning of favours. Very often it was noticed Chinese friends conscientiously trying to outdo each other in terms of favour-giving, something which they said was expected of friends but not strangers. These need not be big or expensive but may include small gestures like buying food for each other, helping each other out in their work, and the sharing of lecture notes. The important thing was that the obligation they put on each other would foster and sustain the relationship. Interestingly, this ‘obligation’ did not connote a negative meaning, in that a person was being forced to do something for another. Rather, it was a desired and willing attempt at reciprocating favours or gifts,

Reciprocity is the key. For example, supper with a friend. If I pay this time, the friend should ‘*zi dong*’ [observe proper behaviour of reciprocation] pay next time. If I keep paying, there is no reciprocity. People shouldn't take advantage of us who try to pay

first. So far I've not had problems. But with Aussies, yes. Like if you insist on paying, they'll actually let you, even though you have paid many times. Not '*kek-ki*' [to stand on ceremony] *lah!* (Singaporean male, No. 33)

To some of these Chinese students, the Australian practice of splitting or sharing the bill seemed quaint, especially when it came to big occasions like birthdays and other celebrations, where the Chinese host was expected to pay. In addition, to the Chinese students, mutual obligations extended beyond the two individuals, to include a social network of relationships with other in-group members. This network or *guanxi* acted like a support system from which one could seek help (for example, getting a job) or enhance one's own status (for example, making friends with one's tutor or lecturer):

Networking is very important. A lot of the things you do need to depend on '*guanxi*', lor. Like getting a job or getting things done. That kind of thing. (Singaporean female, No. 42)

Like whenever I consider helping someone, I will think '*wo gen ni you shen me guanxi?*' [whether I have some sort of "*guanxi*" with you], then I will help you. (Singaporean male, No. 10)

The Chinese students' style of interaction, where highly involved communication was only reserved for their in-group or people whom they had known for some time, may be the reason why many of them found the Australian show of outward affection, even to strangers, puzzling, pretentious, superficial and insincere. Perhaps this was why 73 per cent of Malaysians, 74 per cent of Singaporeans, but no Australians cited **Sincerity** as an important characteristic for successful communication. For many of the Chinese informants, one should have limited or at most, an aloof style of interaction with strangers. In addition, there seemed to be the implication that one should not trust strangers. This was indicated by 73 per cent of Malaysians and 49 per cent of Singaporeans bringing up *Trust* as a key determinant of successful communication, but no Australians did so.

They [the Australians] are more easy going. For us, we take a long time to trust and make friends. (Singaporean male, No. 30)

In Singapore, people on the streets do not smile at you or even sales people in the shops. But like my friend said, when people here ask you how you are, it's like very superficial. (Singaporean female, No.17)

Managing Conflicts and Negotiating

From Table 1, 80 per cent of Malaysians and 49 per cent of Singaporeans cited **Negotiation Styles** as critical for successful encounters while no Australians did so. These students' negotiation styles during conflict situations involved largely the consideration and management of own-face, other-face and mutual-face. Many of the Chinese informants tended to avoid using direct and aggressive methods of conflict resolution. Instead, maintaining harmony seemed to be more important than **honesty** in revealing one's emotions.

[On disagreeing with an assignment group mate] Usually, I'll withhold my thoughts. I can't say anything. I mean I can say, I can do something, but I did not want to do it. And I don't know how to put it. And I don't know if I do something, or say something, it might affect the relationship. So I kept my thoughts. (Malaysian male, No. 02)

The Chinese students displayed unwillingness to confront and argue with others in public. This was evident when some of them were reluctant to confront their Australian neighbours for making noise and disturbing them,

Better not say anything ... They might take revenge on us. (Malaysian male, No. 08)

However, when one of the Chinese groups had a weekend evening gathering session, their Australian neighbours went over at around 11.30pm and told them to quiet down. The Chinese students complied but commented that since “the Aussies don't give us face,” they would do the same next time (that is, when the Australian neighbours had a party, they would go over to tell them to keep the noise down). Nevertheless, they never did on the several occasions when opportunity presented itself.

It's All in the Upbringing

The style of **upbringing** (45% of Australians; 80% of Malaysians; 49% of Singaporeans) was also highlighted as a factor for the students' preferred communication styles. Many of the Australian students, even though they recognised their bond with and dependence on their parents, stressed that they were brought up mainly to be their 'own person'. They were taught to think independently and, if needed, to defend their own views. Not surprisingly, 60 per cent of Australians cited **Having Own Opinions** as a key factor of successful communication compared to only 13 per cent of Malaysians and 8 per cent of Singaporeans. This showed through in their open and unreserved communication with their parents,

There is also not so much fear to disappoint your parents. You are your own person, you don't have to live up too much to their expectations. No doubt you want them to be proud, but not so much an emphasis on academic things. (Australian female, No. 23)

In contrast, the communication behaviour of the Chinese informants with their parents tended to be affected more by their roles (as children) rather than as individuals. All the Malaysians, 96 per cent of Singaporeans but no Australians cited **Role Responsibility** as a criterion for successful communication. Strict guidelines or norms attached to these roles, such as respect for elders and obeying the wishes of elders, in turn accentuate their close association with their parents,

The way we treat our elders is different. Like for them, if they are filial (toward) their parents, they do it because they want to. But for us we might go like “Oh, shit! I've got to look after my mum.” Of course I won't say that out. But we do it because we have to. (Malaysian female, No. 07)

Some of them, like they treat their parents like friends. They can argue with their parents, like nothing wrong. (Singaporean female, No. 06)

I know that we look at our parents differently. I know that a lot of people that I've talked to from Asian background, it is very important what their parents think. (Australian male, No. 21)

“I am afraid what I say might be wrong”

The Chinese students' collectivistic view of the self was also evident in the way they performed in class. During the tutorials, the authors observed that these students were quite reluctant to participate in class discussion. When this observation was brought up during the interviews, the authors found that it was because they expected to get the answer wrong or that other students were better than them (whether at verbal fluency, or at getting the answer right). This problem was further exacerbated by the expectation that they should provide the correct answer. As a result of such self-effacing (0% of Australians, 90% of Malaysians and 61% of Singaporeans) or abasing outlook (0% of Australians, 67% of Malaysians and 49% of Singaporeans), many of them preferred an environment that allowed them to present structured answers, such as in a presentation of a specific topic, rather than engage in open discussion where they might have to voice their opinions,

I don't know what I am saying, whether it'll be correct or is it wrong. I will keep my thoughts to myself if I am not very sure about it... In a way, presentation is good because then you are really, really prepared for that topic. It is fair because you will be the one speaking and everyone will have their turn to speak, you don't have to argue it out in class. (Singaporean female, No. 22)

When I speak up, I am afraid what I say might be wrong. But the Australians don't really think, they just say. (Malaysian male, No. 10)

We are made more to accept than to have our own views. When I was in school, I was told, 'Don't raise your hand unless you have a question' or 'Don't talk unless you are spoken to.' Yeah, they tell you to be critical but according to their rules and form. (Singaporean female, 09)

Unity is power – “More brains are better than one”

The Chinese students' approach to work also encompassed a need for group unity, consensus and interdependence. Applied to group assignments, many of them preferred to work on the whole project collectively rather than assign a section to each member. The assumption was that working together as a group would lessen the likelihood of getting the answer wrong. In addition, no one would be able to mangle. The Australian students, on the other hand, took on a more individualistic and independent approach where each group member was responsible for a topic or section. Some of the Chinese students agreed that the “Aussie way” was more efficient, but more risky as well (in getting it wrong),

It is quite different from my other group projects. Because all along, we will get together and discuss and sometimes if the whole thing stretches for hours, you just go on till everyone is tired. Whereas this one [with Australian students] just go, “You do this part”. With my Singaporean mates, we plough through it together. More brains are better than one. By coming together, everybody has to do his or her fair share. If you separate out and allocate work, what do you do if that person doesn't come back with that work. If you do it collectively, no matter what, you still got to participate in the work or meeting. But I do think that the Aussie way seems more efficient, not that time wasting. But the fact that you are discussing it as a group, chances of you getting it wrong is actually much lower. Because you are actually reasoning until everybody sees the logic. Everyone agrees. But if you divide the work up and I work my part from home, it would seem logical only to me. (Singaporean male, No. 40)

We usually assign a section to each person and let the person go and take charge of that section, chapter, topic. Each person will get to pick what he or she wanted to do. (Australian female, No. 38)

I worked with an Australian last year and this year and I find they generally have very different work ethic. They are more relaxed. Not like us... *'kiasu'* [Being afraid to lose]. We really do things a month ahead. Generally they are very flexible. They can think of things you never thought of. For us, we think in a straight, narrow way. Follow the text, the structure. Very structured, this way means this way. For them they can think lateral, horizontal, and any kind of way. That is good because it promotes variety. They explore a lot of opportunities we don't see. But it is very risky as well. Especially, we are very marks oriented. (Malaysian male, 21)

We try to meet on a weekly basis. And set out tasks to be done for that week. And then review it, meet for an hour or two. But if someone needs to go somewhere, we'll generally say that's as much as we can do. (Australian female, No. 36)

Safer to get Asian friends. *Ang mo* [referring to Caucasians] do assignments very funny one. You do your part, I do my part. Later combine, right? I don't read your part, you also don't read my part. Then submit. But Asian, we have to see that there is consistency everything. *Ang mo* don't care. You simply do your part. Up to you, *lor!* (Malaysian male, 11)

I am lucky to have mostly Asians in my class. So group work no problem. Unity is power. (Singaporean female, No. 01)

DISCUSSION

Consistent with the views of Markus and Kitayama (1991), the communication tactics of the Australian students in this study reflected a more independent self, while the Chinese students tended to reflect an interdependent one. Table 2 provides a summary of the important communication outcomes and methods the students subscribe to.

Table 2. A comparison of the students' key communication outcomes and methods

VALUES	CHINESE STUDENTS		AUSTRALIAN STUDENTS	
	Important communication outcomes	Preferred communication method	Important communication outcomes	Preferred communication method
Individualism/ Independent self versus Collectivism/ Interdependent self	<ul style="list-style-type: none"> • Preservation of group harmony • Maintaining sensitivity to situational context • Anticipating relevant others' needs and feelings • Observing in-group/out-group distinction • Regarding self as extension of others in in-group • Maintenance of self-, other- and mutual-face 	<ul style="list-style-type: none"> • Consider other people's feelings before one acts (e.g. avoid conflict or gain approval by fitting in with others; self-effacing behavior) • Read others' mind • Expect those in the in-group (e.g. friends, family) to share a sense of unity and interdependence • Treat in-group members differently than out-group members (e.g. networking of friends) 	<ul style="list-style-type: none"> • Respect for individual rights to speech and thought • Advocating independent thoughts and actions • Differentiating the self from others • Realizing and exhibiting individual internal characteristics. 	<ul style="list-style-type: none"> • Assert and express individual opinions • Say what's on your mind • Be frank • Show limited concern for what others think of you (e.g. offering friends honest but constructive criticism)

For instance, individuality seemed to be positively valued and there seemed to be a stronger desire among the Australian students to assert their individuality when communicating with others, including parents and friends, as well as in class situations. Although these students also saw the need to affiliate with others, it was more a result of common interests. More often than not, the emphasis in these relationships was on the individual and his or her distinct and unique inner qualities (for example, outgoing, fun loving). Based on the student-interviews and observations, individuality can be achieved through the person's possession, as well as expression of his or her own views. This was especially so when they were passionate about the issues. In addition, their behaviours seemed to be guided more by their individual personalities rather than by their roles (for example, as children or students). A good example was the Australian students' apparent nonchalant attitude during assignment and examination times. Whereas the Chinese students hinted that such behaviour was unbecoming of a model student, the Australian students did not seem too affected by others' opinions of them. In addition, although they wanted their parents to be proud of them, they believed that essentially they were their 'own person.'

In contrast, a high degree of self-monitoring seemed to apply to the communication styles of the Chinese students. These students had a tendency to be very sensitive to the feelings of others and

showed a high inclination to act and react according to others' anticipated perceptions of their behaviour. This was apparent in their reluctance or feeling '*pai-seh*' [feeling shy] to voice their opinions, especially in class or in a group, for fear of imposing their views on others. Some researchers further argued that the Chinese attention to and consideration for others may actually use up precious cognitive space which otherwise might be devoted to individual and creative expression, thus resulting in their lack of verbal fluency (Liu and Hsu 1974; Liu 1986). In addition, the Chinese students were more likely to put themselves down or downplay their abilities (abasement) in relation to others. These findings are not surprising since recent studies have shown that there is a tendency for Chinese people, even in the most rapidly modernising parts of the world, to act in accordance with the anticipated expectations of others instead of with internal wishes or personal characteristics (for example, Yang 1981; Bond 1986). In fact, Bond et al. (1982) found that in Hong Kong college students providing self-effacing or humble attributions following success were liked better by their mates. Indeed, individuals with interdependent selves will generally not assert that they are better than others, will not express satisfaction at being superior to others and truly may not enjoy it. As one of the informants put it,

As long as I know I am good, enough already. We were not brought up to brag about ourselves. (Singaporean male, No. 05)

The interdependent selves of the Chinese students were also revealed through the importance placed on achieving a sense of security by affiliating with a group through similar language or cultural background, as well as their intense need to work collectively as a group and co-operatively. However, it should be noted that an interdependent self does not indicate the absence of the self, including individual judgments, opinions and abilities, or that the self is fused with others (Markus and Kitayama 1991, p.228). Neither does it imply a lack of self-agency. Ironically, it can be seen from the Chinese students' behaviour and responses that it requires a great sense of self-awareness and agency in order to adapt and accommodate oneself to various interpersonal situations.

***Renqing* and *Guanxi*: The Art of Reciprocity**

Two important Chinese concepts, which characterise the relatedness of the interdependent self with others and were brought up by the Chinese students, are *guanxi* and *renqing*. *Guanxi* refers to the status and intensity of an on-going relationship between two individuals that extends to involve others who are part of the social networks of the two individuals (Kirkbride, Tang and Westwood 1991, p.370). *Renqing* essentially refers to the Chinese belief that the reciprocity of conducts and emotions, including favour and hatred, reward and punishment, between human beings, should be as absolute as cause-and-effect. Thus when a Chinese individual acts, a response or return is usually expected (Yang 1957). *Renqing* is attractive in Chinese relationships first because it indicates of the social rules that one has to follow in order to interact successfully with other individuals. Second, it acts as a resource through which a person can present to another a gift, for instance the giving of favours (Hwang 1987; Cheng 1988). In general, there are two social situations where the *renqing* rule may apply: First in daily situations where contact with individuals within one's social network or *guanxi* should be kept through regular greetings, visitations and the exchange of gifts. Second, in times of trouble, when an individual in one's network faces problems, such as needing to get a job, sympathy, help and *renqing* should be offered to that person, which ideally should be returned as soon as possible (Hwang 1987; Cheng 1988).

In this study, the importance placed on observing *renqing* and *guanxi* by the Chinese students was recurrent in the way they viewed the formation of friendship. Many of these students would only form relationships with those with whom they saw the likelihood of interdependency, where material or emotional favours could be exchanged in order to form some kind of emotional obligation (Pye 1982). In addition, before offering a favour or *renqing* to someone, the Chinese

students were more likely to consider the level of *guanxi* they had with that person. Yum (1988) refers to this type of interaction as anticipatory communication because both parties expect some kind of mutual benefit (material or emotional) from the relationship. Such anticipatory communication requires a high degree of social sensitivity because there is an implicit expectation for individuals to anticipate others' needs or to know their feelings without asking or being told, and vice versa; to do otherwise would indicate poor social skills or character deficit (Yum 1988).

Although the concept of reciprocity was also prevalent in the Australian students' friendships, it did not connote the intensity of obligation in the Chinese relationships. Instead, the word 'obligation' may carry a negative connotation with regards to friendship in a Western context. Wierzbicka (1997, p.211) uses the example of the practice of 'shouting a treat' (that is, paying for a treat) to illustrate the Australian sense of reciprocity. She argues that the emphasis on reciprocity in the Australian sense is not strictly equivalent to the repayment of a debt. Instead, the obligation is for the recipient to join in the merriment and share in the companionship. In addition, according to Willcoxson (1992), although Australians can be very group orientated in that they believe that it is appropriate and essential to be responsible and loyal towards one's relevant others such as friends, family and neighbours, individuality and distinctiveness were equally, if not more, valued. Group orientation, if forcibly imposed, will usually be rejected.

The In-group versus Outgroup

Although communication behaviour among Chinese may be very group-oriented, it does not apply to every individual they come into contact with. Instead, a differentiation is often made between those individuals who belong to one's in-group (*zijiren*) and those who are outsiders (*wairen*); attention is only devoted to the needs, desires and goals of *zijiren* (Ho 1988). Indeed, recent studies have found that individuals in collectivistic cultures tend to behave in extremely individualistic ways when dealing with members outside their social network (for example, Furnham and Stringfield 1993; Gabrenya and Hwang 1996). Examples in this study included the Chinese students' outward aloofness and extreme politeness toward people outside their social circle, and their profound need to work with people within their social network.

Previous studies have also found that Chinese individuals revealed more about themselves to those in their in-group than did their Western counterparts (Gudykunst, Gao, Schmidt, Nishida, Bond, Leung, Wang and Barraclough 1992; Goodwin and Tang 1996). Such rigid communication styles not only place people into various relational realms, but also prescribe specific rules for human interactions (Wiemann, Chen and Giles 1986). As a result, many Chinese feel inept at dealing with strangers, and although they will become highly involved in conversations with someone they know, they would rarely speak to strangers (Gao, Ting-Toomey and Gudykunst 1996). This in-group/out-group distinction was apparent in this study through the Chinese students' views that high degrees of warmth and graciousness should only be afforded to friends. Thus, they were puzzled by Australian's general outward display of friendliness toward other people, even strangers. In social relationships, it is common for Chinese individuals to initially exhibit attitudes which indicate distrust towards others, in particular to people who are considered out-group members (Kau and Yang 1991). This is because the Chinese people are constantly assessing other peoples' *cheng ken* (or sincerity) and *xing yong* (or trustworthiness). These values are considered important in any *guanxi* or interpersonal relationships with in-group members. As such, treating strangers or acquaintances in a congenial way would be viewed as devaluing friendship (Yum 1988).

IMPLICATIONS AND CONCLUSIONS

In the world of today, the international success of many countries relies heavily on the ability to create and sustain excellent cross-cultural relationships. However, it would be difficult to achieve

this without good communication. Hence, shortcomings notwithstanding, the results of this study should serve as an impetus for countries, educational institutions, instructors as well as the students themselves to employ different strategies to improve intercultural communication competency and hence intercultural encounters. More importantly, the findings should provide a reference point for understanding cultural differences and cultivating awareness of their nuances and requirements as a basis for communicating.

At the national level, in order to respond to and compete in a globalised world market, understanding and communicating with foreign cultures will be inevitable. In the same vein, many industries are dealing with a multi-cultural work force. In order to work and live harmoniously, these culturally different individuals will need to learn to appreciate and relate with each other. At the school level, the consideration of communication strategies may be helpful in facilitating and increasing participation in existing counselling programs. At the instructor's level, the results of this study advise the rethinking of various instructional methods in the classrooms.

Finally, the understanding, management and accommodation of cultural differences by the students themselves will definitely improve the probability and quality of their intercultural interaction experiences. For instance, the Australian students in this study need to understand that the formal and distant disposition of the Chinese students may not be an indication of aloofness or unfriendliness, but a lack of knowledge in dealing with strangers as a result of the strict rules of communication regarding in-group and out-group members. On the other hand, the Chinese students should appreciate that verbosity, verbal fluency and the ability to express oneself are signs of confidence, not arrogance or cockiness, which are generally esteemed by Australians. Indeed, the impact of international students on their host countries cannot be underestimated. These individuals have the potential to influence the political and economical decision-processes of their home countries. Hence, good relationships between host and international students can translate to better relationships between their countries.

REFERENCES

- Albert, R.D. and Triandis, H.C. (1985). Intercultural education for multicultural societies. *International Journal of International Relations*, 9, 319-337.
- Blight, D. (1997) *Issues in Internationalisation of Australia's Higher Education*. Canberra: IDP.
- Bochner, S., Hutnik, N. and Furnham, A. (1986). The friendship patterns of overseas and host students in an Oxford student residence. *Journal of Social Psychology*, 125, 689-694.
- Bond, M.H. (Ed.) (1986). *The Psychology of the Chinese People*. Hong Kong: Oxford University Press.
- Bond, M.H. (Ed.) (1996). *The Handbook of Chinese Psychology*. Hong Kong: Oxford University Press.
- Bond, M.H., Leung, K. and Wan, K.C. (1982). The social impact of self-effacing attributions: The Chinese case. *Journal of Social Psychology*, 118, 157-166.
- Brislin, R., Lonner, W.J. and Thorndike (1973). *Cross-cultural Research Methods*. New York: John Wiley and Sons.
- Burke, B.D. (1986). *Experiences of Overseas Undergraduate Students*. Sydney: The University of New South Wales Press.
- Cheng, T.C. (1988). A theoretical analysis and empirical research on the psychology of face [in Chinese]. In K.S. Yang (Ed.) *Chinese Psychology*. Taipei: Laureate Book Company, 155-237.
- Chu, G. C. (1985). The changing concept of self in contemporary China. In A.J. Marsella, G. De Vos and F.L.K. Hsu (Eds.) *Culture and Self: Asian and Western Perspectives*. London: Tavistock, 252-277.

- Collier, M.J. and Thomas, M. (1988). Identity in intercultural communication: An interpretive perspective. In Y. Kim and W. B.Gudykunst (Eds.) *Theories of Intercultural Communication*. Newbury Park, Sage Publications, 99-120.
- D'Andrade, R.G. (1992). Schemas and motivation. In R.G. D'Andrade and C. Strauss (Eds.) *Human Motives and Cultural Models*. Cambridge: Cambridge University Press, 23-44.
- Eggerton, L.L. and Langness, L.L. (1974). *Methods and Styles in the Study of Culture*. San Francisco: Chandler and Sharp.
- Fetterman, D.M. (1989). *Ethnography: Step by Step*. Newbury Park: Sage Publications.
- Furnham, A. and Stringfield, P. (1993). Personality and work-performance: Myer-Briggs Type Indicator correlates of managerial performance in two cultures. *Personality and Individual Differences*, 14, 145-153.
- Gabrenya, Jr. W.K. and Hwang K.K. (1996). Chinese social interaction: Harmony and hierarchy on the good earth. In M.H. Bond (Ed.) *The Handbook of Chinese Psychology*. Hong Kong: Oxford University Press, 309-321.
- Gao, G. (1998). An initial analysis of the effects of face and concern for "other" in Chinese interpersonal communication. *International Journal of Intercultural Relations*, 22(4), 467-482.
- Gao, G., Ting-Toomey, S. and Gudykunst, W.B. (1996). Chinese Communication Process. In M.H. Bond (Ed.) *The Handbook of Chinese Psychology*. Hong Kong: Oxford University Press, 280-293.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.
- Goodwin, R. and Tang, C.S.K. (1996). Chinese personal relationships. In M. H. Bond, (Ed.) *The Handbook of Chinese Psychology*. Hong Kong: Oxford University Press, 295-308.
- Gudykunst, W.B., Gao, G., Schmidt, K.L., Nishida, T., Bond, M.H., Leung, K., Wang, G. and Barraclough, R. (1992). The influence of individualism-collectivism on communication in ingroup and outgroup relationships. *Journal of Cross-cultural Psychology*, 23, 196-213.
- Hall, E.T. (1977). *Beyond Culture*. New York: Anchor Doubleday.
- Hecht, M.L., Collier, M.J. and Ribeau, S. (1993). *African American Communication: Ethnic Identity and Cultural Interpretation*. Newbury Park: Sage Publication.
- Ho, D.Y.F. (1988). Cognitive socialization in Confucian heritage cultures. In P.M. Greenfield and R. R. Cocking (Eds.) *Cross-Cultural Roots of Minority Child Development*. New Jersey: Lawrence Erlbaum, 285-313.
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-related Values*. Beverly Hills: Sage Publications.
- Hofstede, G. and McCrae, R.R. (2004). Personality and Culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research*, 38(1), 52-88.
- Hsu, F.L.K. (1985). The self in cross-cultural perspective. In A.J. Marsella, G. De Vos; and F.L.K. Hsu (Eds.) *Culture and Self: Asian and Western Perspectives*. London: Tavistock, 24-55.
- Hwang, K.K. (1987). Face and favour: The Chinese power game. *American Journal of Sociology*, 92(4), 944-974.
- Kau, A.K. and Yang, C. (1991). *Values and Lifestyles of Singaporeans: A Marketing Perspectives*. Singapore: National University of Singapore Press.
- King, A.Y. and Bond, M.H. (1985). The Confucian paradigm of man: A sociological view. In W.S. Tseng and D.H. Wu (Eds.) *Chinese Culture and Mental Health*. Orlando: Academic Press, 29-45.
- Kirkbride, P.S., Tang, S.F.Y., and Westwood, R.I. (1991). Chinese conflict preferences and negotiating behavior: Cultural and psychological influences. *Organization Studies*, 12(3), 365-386.
- Kluckhohn, C. and Strodtbeck, F.L. (1961). *Variations in value orientations*. Evanston: Harper, Row and Peterson.

- Liu, I.M. (1986). Chinese cognition. In M.H. Bond (Ed.) *The Psychology of the Chinese People*. Hong Kong: Oxford University Press, 73-105.
- Liu, T.Y. and Hsu, M. (1974). Measuring creative thinking in Taiwan by the Torrance test. *Testing and Guidance*, 2, 108-109.
- Markus, H.R. and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.
- Mirande, A. and Tanno, D.V. (1993). Understanding interethnic communication and research: 'A rose any other name would smell as sweet.' *International Journal of Intercultural Communication*, 17, 381-388.
- Mullins, G.; Quintrell, N.; and Hancock, L. (1995). The experiences of international and local students at three Australian universities. *Higher Education Research and Development*, 14(2), 201-231.
- Munro, D. (1997). Levels and processes in motivation and culture. In D. Munro, J.F. Schumaker and S. C. Carr (Eds.) *Motivation and Culture*. New York: Routledge, 3-15.
- National Differences. In U. Kim, H. Triandis, C. Kagitcibasi, S.C. Choi and G. Yoon (Eds.) *Individualism and Collectivism: Theory, Method and Applications*. Thousand Oaks: Sage Publication, 85-119.
- Nesdale, D., Simkin, K., Sang, D., Burke, B. and Fraser, S. (1995). *International Students and Immigration*. Canberra: AGPS.
- Nowak, R.; Weiland, R. and McKenna, K. (1998). Cultural adjustment in cross-cultural crossing: The function of in-groups and out-groups. Paper presented in the 12th Australian Education Conference. Canberra, Australia: IDP.
- Porter, R.E. and Samovar, L.A. (1988). Approaching intercultural communication. In L.A. Samovar and R.E. Porter (Eds.) *Intercultural Communication: A Reader*, 5th ed. Belmont: Wadsworth, 14-30.
- Pye, L.W. (1982). *Chinese commercial negotiation style*. New York: Oelgeschlager, Gunn and Hain.
- Rokeach, M. (1968). *Beliefs, Attitudes and Values: A Theory of Organization and Change*. San Francisco: Vossey-Bass.
- Samovar, L.A. and Porter, R.E. (Eds.) (1991). *Intercultural Communication: A Reader*, 6th ed. Belmont: Wadsworth.
- Schneider, D. (1976). Notes toward a theory of culture. In K. Basso and H. Selbey (Eds.) *Meaning in Anthropology*. Albuquerque: University of New Mexico Press, 197-220.
- Schram, J. and Lauver, p.(1988). Alienation in international students. *Journal of College Student Development*, 29, 146-150.
- Schwartz, S. (1994). Cultural dimensions of values: Toward an understanding of national differences. In U. Kam, H.C. Triandis, C. Kagitcibasi, S.C. Choi, G.Yoon, (Eds.), *Individualism and Collectivism: Theory, Method, and Application* (pp. 85-119). Sage: Newbury Park, CA.
- Schwartz, S.H. and Sagiv, L. (1995). Identifying culture-specifics in the content and structure of values. *Journal of Cross-Cultural Psychology*, 26, 92-116.
- Shweder, R.A. and Bourne, E.J. (1984). Does the concept of the person vary cross-culturally? In R.A. Shweder and R.A. Levine (Eds.) *Culture Theory: Essays on Mind, Self, and Emotion*. Cambridge: Cambridge University Press, 158-199.
- Tanaka, T., Takai, J., Kohyama, T., Fujihara, T., and Minami, H. (1997). Effects of social networks on cross-cultural adjustment. *Japanese Psychological Research*, 39(1), 12-24
- Todd, p.and Nesdale, D. (1997). Promoting intercultural contact between Australian and international university students. *Journal of Higher Education Policy and Management*, 19(1), 61-76
- Triandis, H.C. (1995). *Individualism and Collectivism*. Boulder: Westview Press.

- Vogt, E.Z. and Albert, E.M. (1970). The 'Comparative study of values in five cultures' project. In E.Z. Vogt and E.M. Albert (Eds.) *People of Rimrock: A Study of Values in Five Cultures*. New York: Atheneum, 1-33.
- Volet, S. E. and Ang, G. (1998). Culturally mixed groups on international campuses: an opportunity for inter-cultural learning. *Higher Education Research and Development*, 17(1), 5-23.
- Volet, S. E. and Pears, H. (1994). *International students in Technical and Further Education (TAFE) Colleges in Western Australia: Students' Reflections on their Experiences and Perceptions of the Future Associations between their Country and Australia*. Murdoch University and TAFE International (WA) Research Report (Perth, CCTN).
- Volet, S. E. and Pears, H. (1995). *ELICOS Students: Reflections on Studying English at TAFE Western Australia*. Murdoch University and TAE International (WA) Research Report (Perth, CCTN).
- Volet, S.E. and Renshaw, P.D. (1993). Cross-cultural differences in university students' goals and perceptions of study settings for achieving their own goals. Paper presented at the Annual Meeting of the Australian Association Research in Education Symposium: Overseas students in Australian Universities – Comparative and Longitudinal Studies of Students' Expectations, Goals and Interactions. Perth, Western
- Volet, S.E.; Renshaw, P.D. and Tietzel, K. (1994). A short-term longitudinal investigation of cross-cultural differences in study approaches using Biggs SPQ questionnaire. *British Journal of Educational Psychology*, 64, 301-318.
- Westwood, M.J. and Bergen, W.A. (1988). A culturally embedded model for effective intercultural communication. *International Journal for the Advancement of Counseling*, 11, 115-125.
- Wiemann, J., Chen, V. and Giles, H. (1986). Beliefs about talk and silence in a cultural context. Paper presented to a meeting of the *Speech Communication Association*. Chicago, November.
- Wierzbicka, A. (1997). Australian key words and core cultural values. In A. Wierzbicka (Ed.) *Understanding Cultures through their Key Words: English, Russian, Polish, German and Japanese*. NY: Oxford University Press, 198-234.
- Willcoxson, L. (Ed.) (1992). *Asia is your Business*. Research Institute for Asia and the Pacific: University of Sydney.
- Yang, K.S. (1981). Social orientation and individual modernity among Chinese students in Taiwan. *Journal of Social Psychology*, 113, 159-170.
- Yang, K.S. (1992). Do traditional and modern values coexist in a modern Chinese society? [in Chinese]. In *Proceedings of the Conference on Chinese Perspectives on Values*, Taipei: Center for Sinological Studies, 117-158.
- Yang, L.S. (1957). The concept of pao as a basis for social relations in China. In J.K. Fairbank (Ed.) *Chinese Thought and Institutions*. Chicago: University of Chicago Press, 291-309.
- Yum, J.O. (1988). The impact of Confucianism on interpersonal relationships and communication patterns in East Asia. *Communication Monographs*, 55, 374- 388.

Comparing university academic performances of HSC students at the three art-based faculties

Noor Azina Ismail

University of Malaya, Malaysia nazina@um.edu.my

Azmah Othman

University of Malaya, Malaysia g3azmah@um.edu.my

University Malaya enrolls students from all states in Malaysia as well as a small number of students from overseas. The objective of this paper is to investigate the effect of past performance on students at three faculties, namely, Faculty of Economics and Administration (FEA), Faculty of Business and Accounting (FBA) and Faculty of Arts and Social Sciences (FASS). Students' prior achievements include their entry scores or points in English language proficiency and mathematics at the Malaysian Certificate of Education (MCE) level. Other factors taken into consideration are gender of students and their ethnic origins. Research results show that entry points are an important factor in influencing students' achievement in all three faculties. Apart from this, female students are found to have better results than their male counterparts in FBA and FAASS. It was also found that mathematics performance at the MCE level is one of the influential factors for academic achievement in FBA.

Academic performance, prior achievement, entry points, language proficiency, mathematics performance

INTRODUCTION

All states in Malaysia have the same education system and ways of measuring achievement and presenting results at the end of secondary school. Lower secondary education lasts three years. Students take a national public examination at the end of this level after which students proceed to a two-year upper secondary education program. At this level, students can either choose one of the three streams of studies: academic, technical or vocational. At the end of upper secondary education, students take another national public examination known as the Malaysian Certificate of Education (MCE). Upon completing upper secondary education, there are several routes to higher education in the country. One of them is a two-year sixth form education program, where students take a centralised examination known as the Malaysian Higher School Certificate (HSC). An alternative route is a pre-university or matriculation program. There are other routes to local universities but those two routes mentioned above are the most common ones. Each route has a separate education system, with different curricula, different approaches to teaching and different ways of measuring achievement and presenting results. Like the sixth form education, the pre-university or matriculation education program also have centralised examinations.

This paper investigates the university academic performance of students from sixth form education in the Faculty of Economics and Administration (FEA), Faculty of Arts and Social Sciences (FASS) and Faculty of Business and Accountancy (FBA) at the University of Malaya, Kuala Lumpur, Malaysia, by examining their prior achievement (university academic entry

points, grades in MCE Mathematics and MCE English), gender and ethnicity. These faculties are assumed to receive students from similar academic backgrounds.

ACADEMIC ACHIEVEMENT

According to Walberg (1984) and Fraser (1987), factors influencing academic achievement can be grouped into three categories. They are student aptitude, instruction and environment. This study is limited by the range of available data and therefore only examines the effect of prior achievement and socio-demographic variables on academic performance in the university examinations. The instructional factors, such as the quantity and quality of education, and the environmental variables such as home environment, university or classroom environment, peer group environment, outside university environment and mass media environment are not investigated.

Academic performance is measured using the Cumulative Grade Point Average (CGPA). Other characteristics and educational outcomes such as leadership, exposure, creativity and motivation are not considered in this study. It is important not to compare directly the CGPA of students from one faculty to another, as this only tells whether the students from one faculty are 'better' or 'worse' than the students from the other. In actual fact, it is a result of different curricula, different approaches of teaching and different ways of measuring achievement. This makes comparisons between faculties difficult. However, comparisons are carried out according to different background characteristics of the students and also different levels of preparedness for tertiary education of the students in each faculty.

Age of students is not taken into account because almost all students are from mainstream post-high school entry and hence none of the students can be considered as mature-aged students. Furthermore a few studies (Archer et al., 1999; Hayes et al., 1997) found older students perform at a level equal to younger students.

Although Leong et al. (1990), in his study of performance of students in three major Malaysian public examinations found that students from urban schools performed better than those from rural schools, studies (Jahara, 2001; Tho, 1994, 1999) carried out in the universities did not find urban-rural status to be a significant factors influencing university performance. Hence these two variables are not included in this study.

On the other hand, ethnicity may play a role in influencing the academic performance in these faculties. In 2001, the Malaysian Prime Minister released the information Malay students were lower, in terms of academic performance when compared to their non-Malay counterparts (*The Sunday Star*, Nation, July 29, 2001, p.2).

The three faculties selected are preferred by the *Bumiputera* (Malays and native people). According to Chew et al (1995), *Bumiputera* students are more inclined to study humanities, social and behavioural science, and commerce and business studies.

Another socio-demographic variable considered in the study is gender. One of the local newspapers (*Berita Harian*, 18 July 2002, p.10) claimed that more than 70 per cent of the students enrolled in the public universities in the academic year of 2000/2001 are female. It was also stated that female students not only perform better at the primary and secondary school examinations but in the university level as well. This was said to be due to a number of factors such as female students were more responsible, serious and showed a higher level of commitment to their studies than did male students.

Apart from the socio-demographic variables, prior academic achievement is also considered as one of the predictors of academic achievement. Though there are many factors that influence academic performance of students, prior academic achievement is taken as a very important

selection criterion at various levels of education in Malaysia. The main determinant for admission to the local public universities is based on university entry points. The entry points can be divided into two sections: points for co-curriculum activities and points for academic achievement. In this study consideration is only given to points for academic achievement to represent prior academic achievement. In addition, the quota system also influenced the admission of students to the university. The quota system has been implemented since 1970s under the New Economic Policy to obtain a better balance in enrolment in the local universities among various ethnic groups. This has resulted in a gradual increase in *Bumiputera* students in local universities.

In this article, the investigation considers the influence of prior English proficiencies and mathematical abilities on academic performance at the university level. Nagaraj and Lee (1992) found that previous mathematical ability was an important factor contributing to academic success in the Faculty of Economics and Administration. This current study, however, only takes into consideration MCE mathematics since only at this level, is mathematics a compulsory subject. English is also an important subject and is included as one of the factors in this study since most of the reference works and textbooks at the university level are written in English. Students with high proficiency in English can be expected to perform better because of this advantage.

THE DATA

Students who were registered in the academic year of 2000/2001 in the three faculties were chosen for this study. The CGPA at the end of the second semester of 2000/2001, ethnic group and gender for all students were obtained from the Centre of Information Technology, University of Malaya. Nonetheless, prior achievements such as university entry points, English proficiency and mathematical ability in the MCE examination, were collected manually from the data held in the University of Malaya Admission and Record Section. The university entry points which consisted of academic and co-curriculum points, had a total of 100 points. The maximum points for academic and co-curriculum were 90 and 10 points, respectively. Only academic entry points were taken into consideration in this study.

Since this study only takes into consideration HSC school leavers, it is appropriate to examine their numbers against the total number of students enrolled in these three faculties. More than 90 per cent of students from FEA and FASS were HSC school leavers. On the other hand, more than half (52 per cent) of students from FBA were from matriculation program, diploma and other courses. Furthermore, since the data on prior academic performance were collected manually from the Admission and Record Section, some of the information was not complete. However, it was possible to collect information on more than 90 per cent of the total students enrolled and under consideration in this study.

Table 1 gives a summary of characteristics of the students enrolled in the second semester of the academic year 2000/2001. More than two-thirds of the students were female. Ethnic origins are divided into two groups: 1 = Malays, 0 = Others/Non-Malays. Among the three faculties, FASS has the largest percentage of Malays (slightly more than half), compared with less than one-third in FEA and FBA. It is expected that better grades in MCE mathematics and MCE English lead to better performance in these faculties. The grades for these two subjects range from 1 to 9 with 1 as the highest grade and 9 as the lowest. These grades are regrouped into two: 1 to 3 as 'excellent performance' and 4 to 9 as 'not excellent performance'. From the table below, most of the students performed very well in MCE mathematics especially students from FBA, where more than 80 per cent of the students obtain grades 1 to 3. However, less than 50 per cent of students from FASS are in this category. Although the percentages of students who do well in MCE English are not as high as that in mathematics, the highest percentage is still from FBA.

Table 1: Characteristics of Students

Variable	Faculty of Economics and Administration	Faculty of Business and Accountancy	Faculty of Arts and Social Sciences
Gender			
Male (%)	31.4	28.4	26.2
Female (%)	68.6	71.6	73.8
Total	204	261	875
Ethnicity			
Malays (%)	37.7	20.7	51.2
Others (%)	62.3	79.3	48.8
Total	204	261	875
Prior Mathematics Performance			
Excellent (%)	58.8	80.8	41.3
Not Excellent (%)	41.2	19.2	58.7
Total	194	260	813
Prior English Performance			
Excellent (%)	35.8	49.6	23.5
Not Excellent (%)	64.2	50.4	76.5
Total	151	258	671

PRIOR AND CURRENT ACADEMIC ACHIEVEMENT

As mentioned above, the CGPAs of students cannot be compared directly from one faculty to another. However, it is still interesting to see the distribution of CGPA across faculties. From Figure 1, it can be seen that the highest and the lowest values of CGPA for FEA and FBA are similar but the box is higher for FBA, indicating that more students in FBA have high CGPA points, when compared to students from FEA. From the figure, it is also evident that the spread of points are different for the three faculties. The CGPAs for students in FBA are wide but for students from FASS, their CGPAs are relatively constant. It is also possible to observe that there are a few outliers in FASS. However, for the purpose of our analysis, these cases are still included.

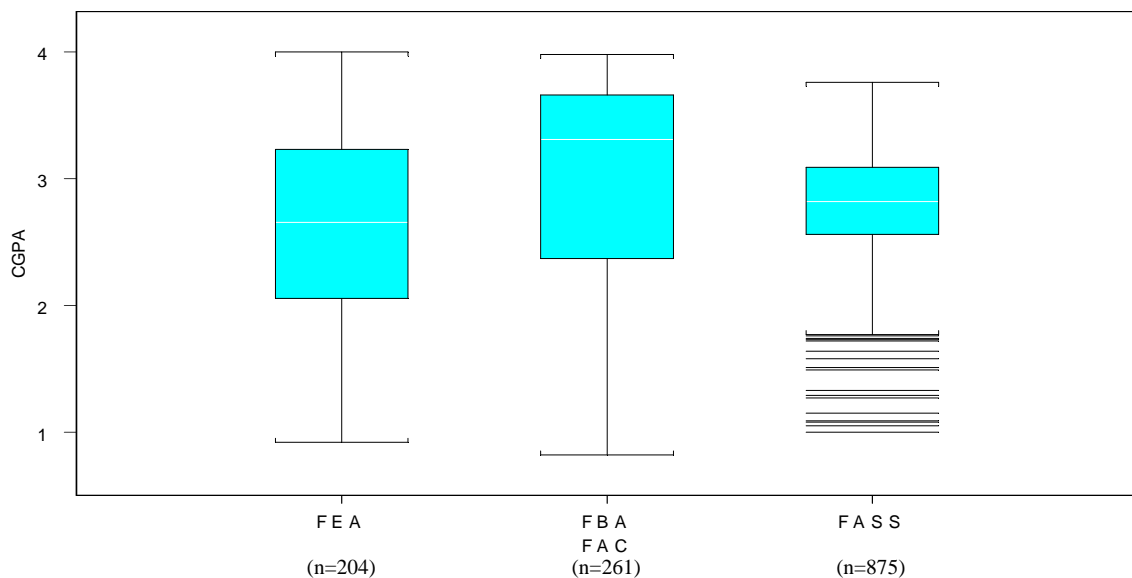


Figure 1. Box-plots for CGPA of students in the three faculties

Examining in greater detail the correlations of the university academic entry points and CGPA, the value of the correlation coefficients are 0.82, 0.82 and 0.26 for FEA, FBA and FASS, respectively. All of the correlation coefficients are significantly different from zero at the 5 per cent level of significance. Although the value of correlation coefficient for FASS is small,

indicating that the relationship is weak. The scatter plots, shown in Figure 2, suggest that students who perform well at the pre-university level will also perform well at the university.

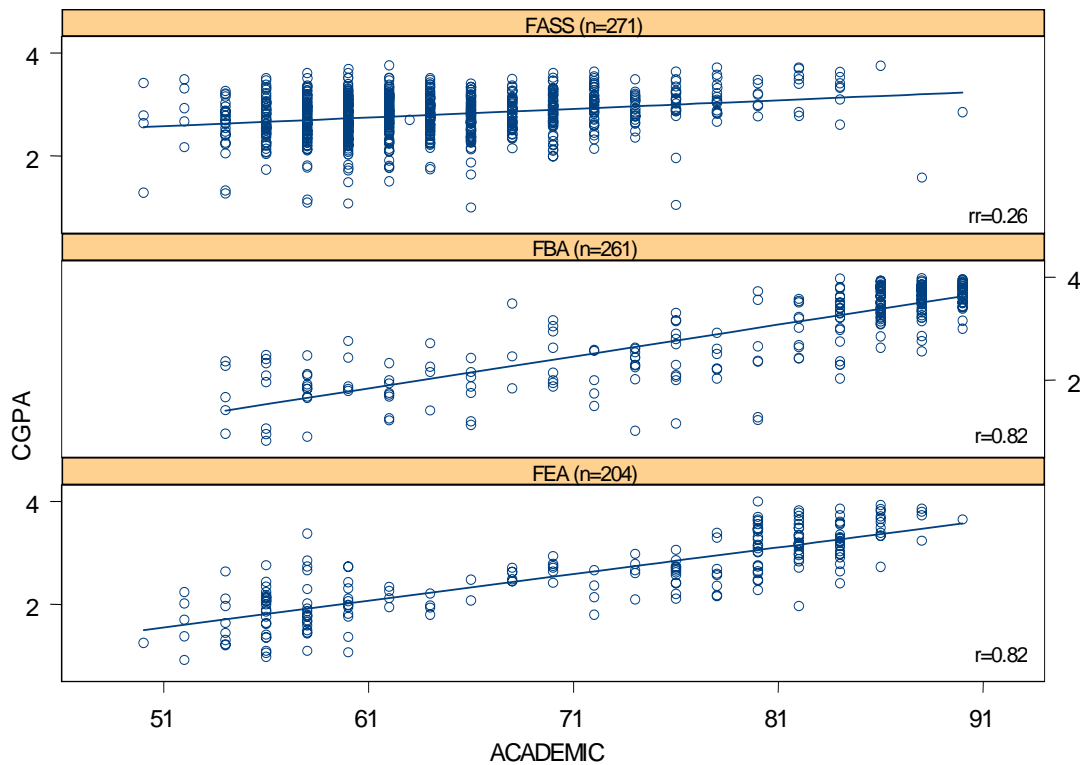


Figure 2. Scatter Plots of Academic Entry Points versus CGPA

Figure 3 shows that there are considerable differences in the entry standard among Malays and non-Malay students in FEA and FBA. This also results in obvious differences in academic success between them. The entry points and the academic performance of Malay students are significantly lower than their non-Malay counterparts. This finding agrees with a similar study by Yahaya et al (2001) on the academic performance of FEA students in the 1998 and 1999 cohort. They found that the academic performance of the Malay students was significantly lower than the Indians and the Chinese. In contrast, although this figure suggests that the differential in entry points and academic performance among the ethnic group is not so obvious for students in FASS, there are recognisable differences in the entry points and academic performance between the ethnic groups.

It is also of interest to analyse student entry points and academic performance by gender. Although the differential in entry points and academic performance for female and male students are not as great as is shown in Figure 4, the figure also indicates that females perform better than their male counterparts. However, the use of significance tests on the data recorded in Figures 3 and 4 are not meaningful since there are likely to be interaction effects between gender and ethnic groups, and academic performance.

REGRESSION ANALYSIS

In this study, we examine five meaningful and potential predictors for explaining academic performance. The regression function has CGPA as the dependent variable. The independent variables are students' academic entry points (ACADEMIC), prior Mathematics performance (MATH, 1 = excellent, 0 = otherwise) and prior English performance (ENG, 1 = excellent, 0 = otherwise), gender (GENDER, 1 = male, 0 = female), ethnicity (MALAY, 1 = Malay, 0 = others). Except for ACADEMIC, other independent variables are binary.

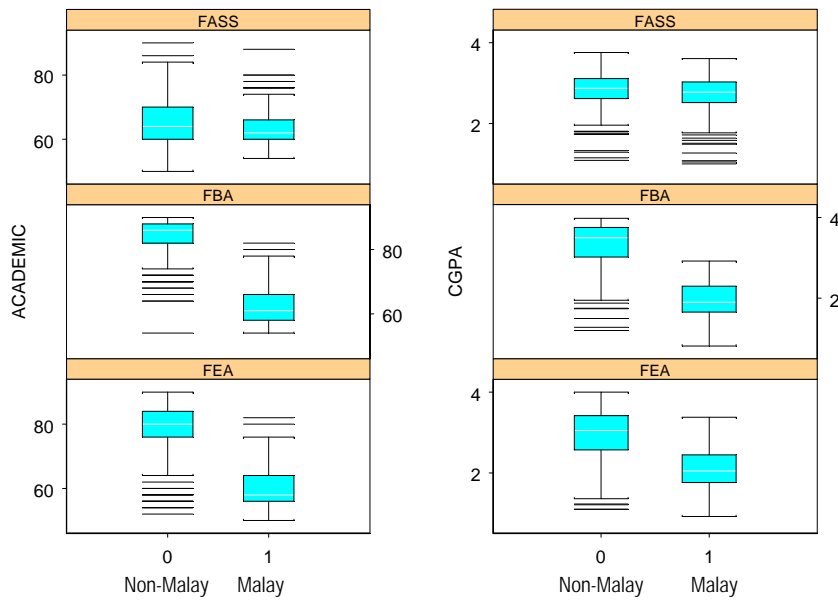


Figure 3. Box-plots for entry points and academic performance according to ethnic groups

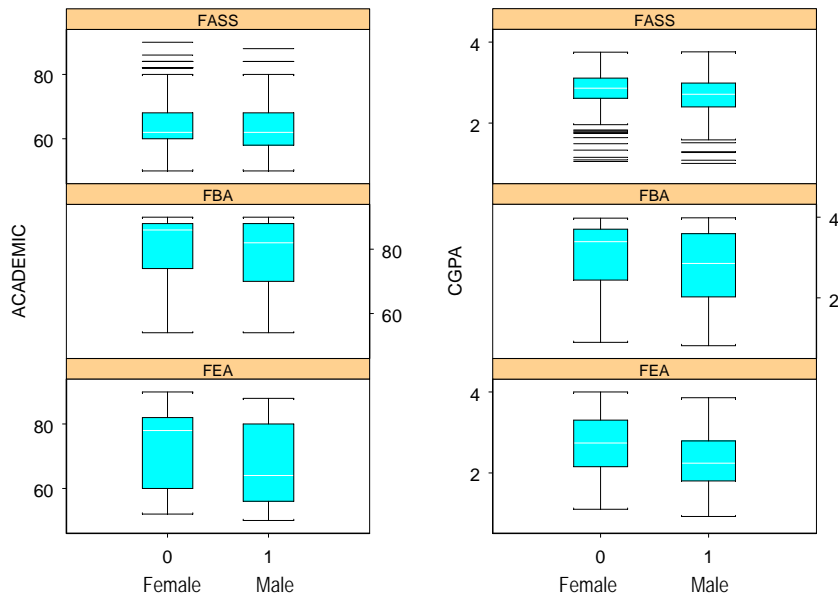


Figure 4. Box-plots for entry points and academic performance according to ethnic groups

Using the multiple linear regression analysis and applying the two potentially competing principles of Best Fit and Parsimony in the data analysis, we find the following results in which only statistically significant effects are recorded.

Model 1: Faculty of Economics and Administration
 $CGPA = -1.09 + 0.05ACADEMIC$

Equation (1)
 Adjusted R-squared = 0.61

Model 2: Faculty of Business and Accountancy
 $CGPA = -1.38 + 0.05ACADEMIC - 0.17GENDER + 0.45MATH$

Equation (2)
 Adjusted R-squared = 0.71

Model 3: Faculty of Arts and Social Sciences
 $CGPA = 1.81 + 0.02ACADEMIC - 0.15GENDER$

Equation (3)
 Adjusted R-squared = 0.10

The regression equations (1), (2) and (3) indicate that, the student performances in the FEA, FBA and FASS are very much depended on the pre-university qualification.

In FEA, the only significant factor is the entry points (ACADEMIC) and the additional four other factors are not significant. ACADEMIC alone explains around 61 per cent of the variance in CGPA in this study. The coefficient of ACADEMIC is positive, indicating that students who perform well in the pre-university qualification also do well in the first year university examination. In the FASS data both ACADEMIC and GENDER are important factors in influencing the academic performance of FASS students. The coefficient of GENDER is negative, indicating that female students perform better than their male counterparts are ACADEMIC points are taken into accounts. The variance explained in the performance of the FASS students is quiet low, around 10 per cent, indicating that other factors related to instruction and environment, that are not considered in this study may have influenced performance. Alternatively, the grading of performance at the university level in this field may be low reliability, thus reducing the explanatory power of the analysis.

For FBA students, ACADEMIC, GENDER and MATH play important roles in influencing CGPA. The coefficient of ACADEMIC and MATH are positive, suggesting that students who perform well during HSC and do well in Mathematics subject during MCE also perform well at the university. The negative coefficient for GENDER indicates that female students again perform better than male counterparts. These three variables explain around 71 per cent of the variation in CGPA.

CONCLUSION

This study finds that the prior achievement in HSC is a reasonable good predictor to first year university performance, for all the three faculties investigated. Besides prior achievement, gender plays an important role in influencing success in the university at least in two faculties, Faculty of Business and Accountancy and Faculty of Arts and Social Sciences. Generally, female students outperformed male students. This can also be due to high prior academic ability among them as compared to the male students. Good ability in Mathematics also helps to influence success in academic performance for students in the Faculty of Business and Accountancy but not for students from the other two faculties. Unlike Faculty of Economics and Administration and Faculty of Business and Accountancy, factors considered in this study have a very low explanatory effect on the academic performance in the Faculty of Arts and Social Sciences. This study also suggests that the problem of differences in academic performance across ethnic groups and gender may operate prior to the university level.

REFERENCES

- Archer, J., Cantwell, R. and Bourke, S. (1999). Coping at university: an examination achievement, motivation, self-regulation, confidence, and method of entry. *Higher Education Research and Development*, 18(1), 31 – 54.
- Fraser, B. J., Walberg, H. J., Welch, W. W. and Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11, 145 – 252.
- Hayes, K., King, E. and Richardson, J. (1997). Mature students in higher education: III. Approaches to studying in access students. *Studies in Higher Education*, 22(1), 19 – 31.
- Leong, Y. C., Cheong, S. Y., Chew, S. B., Salim, S., Marimuthu, T., Talib, A. A., Rashid, A. R. and Chew, S. M. (1990). Factors influencing the academic achievement of students in Malaysian school. Report prepared for the Ministry of Education, in conjunction with the 8th World Bank Loan, Government of Malaysia.

- Nagaraj, S. and Lee K. H. (1992). The determinants of academic performance: A comparison of three cohorts at the Faculty of Economics and Administration, University of Malaya. *Malaysian Journal of Economic Studies*, 29 (2), 47 – 61.
- Tho, L. M. (1994). Some evidence on the determinants of student performance in the University of Malaya introductory accounting course. *Accounting Education*, 3(4), 331 – 340.
- Tho, L. M. (1999). Predicting success in university accounting examination performance. *Journal of Educational Research*, 20, 95 – 104.
- Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership*, 41(8), 19 – 27.
- Yahaya, J., Tey, N. P. and Liaw, S. H. (2001). Chasing the Grades – Academic Performance of First-Year FEA Students. *FEA Research Update*, 1 (2), 2 – 6.

Private education as a policy tool in Turkey

Mustafa Cinoglu

University of Illinois at Urbana-Champaign cinoglu@uiuc.edu

This paper discusses privatisation as policy tool to solve educational problems in Turkey. Turkey, as a developing country, is faced with many problems in education. Large class size, low enrolment rate, girl's education, high illiteracy rate, religious education, textbooks, curriculum and multicultural education are some of the important problems. On the other hand, cultural diversity, economical problems and lack of resources, migration, central management, the rapid growth in school-age population and regional differences make the solution difficult. In principle, privatisation could help in the solution of the problems in Turkey. Private schools can increase the quality of education. Private schools can share responsibility with the government by investing money in education. In addition, private schools can also meet the needs for diversity in society promoting specific religious or philosophical views by using their own teaching methods. The paper describes the availability of private schooling in Ankara as a baseline for thinking about the development of private schooling in Turkey.

Turkey, Ankara, privatisation, policy analysis, private schools

INTRODUCTION

Why do people pay fees for private education instead of using free public schools? Private schools are seen to provide better quality education with greater safety. From a policy perspective, they can meet needs of diversity in society. Private education can also help governments by contributing to educational costs. The importance of educational quality increases because of the relationship between the market economy and education. Parents demand quality education for their children and public schools are seen as failing to provide it: Public schools are bound by rules and regulations; they are managed by government bureaucracies; they cannot choose their teachers and curricula. Teachers are unionised and they have permanent status in public schools. These conditions are seen as reducing the success of public schools. On the other hand, private schools can choose their principals, and teachers. Private schools must provide quality education to keep their students because parents as customers always monitor and judge the success of private schools.

The Turkish educational system has problems of cultural diversity. The highly intensive central management of Turkish education does not meet the needs of a diverse society. Public schools serves the public interest by preparing students for society and by teaching common values, but private schools can better meet private interests by providing alternatives in order to address diversity so that parents can act in the best way according to their perception of the interests of their children.

Private education can also decrease educational costs for the governments. The Turkish Educational Ministry controls the largest component of the state budget, but does not have funds to provide quality education for all students. Moreover, the rapidly increasing population makes it very difficult to solve the educational problems in Turkey. The private sector can invest capital and run schools efficiently.

This paper discusses privatisation as a policy tool to solve the educational problems in Turkey. The question in this paper is “Should privatisation be a policy tool to address educational problems in Turkey?”. The paper has five sections. First, the paper provides information about Turkey and the Turkish education system. Second, it explores privatisation as a new policy tool. Third, it considers private education and its problems. Fourth, the paper answers the question, How should private schools be supported? Finally, the paper discusses private schooling in Ankara and asks what does it involve?

TURKEY AND TURKISH EDUCATION SYSTEM

Turkey is a country which is located in Europe, Asia, and the Middle East. The capital city of Turkey is Ankara and the population of Turkey is 67.8 m. (as of 2002). Ninety-nine per cent of the population of Turkey is Muslim. After September 11, the importance of Turkey increased because Turkey is the only country in which Islam is the major religious community, but is constitutionally secular guaranteeing complete freedom of worship to non-Muslims. In addition, the present conservative government, which has Islamic origins, is engaged in reform, with the goal of Turkey joining of the European Union (EU).

Westernisation policies have a long history in Turkey, beginning in the last period of Ottoman Empire. Modern Turkey was founded in 1923 by Mustafa Kemal Ataturk and the new republic turned to Europe and accepted Western values. Ataturk created the necessary conditions for the modernisation of Turkey with his radical reforms. Turkey applied to join the European Union (EU) in 1964 and, in December, 2004, the EU decided to commence discussion about the membership of Turkey, but full membership of the EU looks uncertain prior to 2015.

The Turkish educational system has to address many problems as a requirement of ‘Westernisation’, including those associated with highly central management, economical and cultural dispersion, diversity necessities, religious education, regional differences, educational quality, class size, enrolment rate, and girl’s education.

One of the major problems in Turkish education is its central management. The Turkish education system has highly centralised governance. However, there are many difficulties associated with governing a large education system from the centre. Central management prevents the efficient allocation of educational resources to local needs. Although the government started to decentralise in 2004 and many of the Ministries (except for the National Defence, Economy, and Education Ministries) delegated their authority to local management.

Religious education is another major issue in Turkey and the scarf is not allowed in the schools. Some parts of the society see the scarf is a symbol of an Islamic Regime, which challenges secular ideals. Others claim that the scarf is a requirement of their religious faith, and wearing it is a democratic right. Many women cannot continue to universities because of the scarf problem. Consequently, if they have enough money, many families send their children abroad for university education. The Prime Ministry, Tayyip Erdogan, said that he had two girls and they were going to universities in the United States because of the scarf problem. This rule decreases the enrolment rates of girls in Turkey. With membership of the European Union (EU), Turkey must consider the problems of religious education. Private education can help to solve this problem by offering extensive educational opportunities for religious groups.

Turkey is a diverse country. Kurdish, Laz, Cerkez and Abaza are the major subcultures in Turkey. In 2004, the Turkish government gave greater rights to diverse cultural groups: The state-run radio and television network, TRT, started to broadcast programs as Kurdish, Arabic, Cerkez and the government gave permission in Kurdish courses. It was assumed that the Kurdish population would request greater educational rights and, under the Westernisation policy, Turkey would have to solve these educational problems.

The status of women is another issue. Some women wear European-style clothes, work outside their home, go to hotels, discos, and beach resorts, and share housework with their husbands. Some women use the scarf traditionally and obey their husbands' or fathers' rules. They spend most of their time on housework and taking care of their children. Other women are conservative Muslims and wear gowns, but their numbers are relatively low. According to the constitution, males and females have equal rights and gender discrimination is prohibited. There has been one women prime minister, many women ministers, and many women in the National Council.

Among OECD countries, Turkey has the highest percentage of females not in education or not in work. The percentage of 15 to 19 year-old females, who are not in education or work, is 48 per cent while the male percentage is 24 per cent (Education at a glance, 2003). The rate of participation to in the labour force in Turkey is 24 per cent female and 75 per cent male (State Institute of Statistics Prime Ministry Republic of Turkey, 2001).

Another major problem in Turkey is regional inequalities, with significant differences between the east and west sides of the country. In the dispersion of Gross Domestic Product (GDP) among the regions, East Anatolia, Southeast Anatolia, and Black Sea are the poorest regions of in Turkey. While East Anatolia had 4.15 per cent of GDP in 1997, Marmara (West side of the country) has 38.14 per cent of GDP (See Table 1) (Gezici, F., Hewings, J.D. 2002). The majority of the population in these regions generally worked in agriculture and stockbreeding; since there was little industrial works. The private sector tended to avoid investment in the eastern part of the country because of transportation problems and terror, and most government employees did not want to work there because of the lack of social facilities, safety and difficult working conditions.

Table 1. Share of gross domestic product (GDP) among geographical regions

Geographical Regions	1980	1985	1990	1995	1997
Marmara	31.76	34.85	35.86	36.74	38.14
Aegean	16.08	16.53	16.48	17.15	16.75
Mediterranean	11.92	10.95	11.29	11.25	11.19
Central Anatolia	18.67	17.81	16.95	16.75	16.06
Black Sea	10.69	9.37	8.97	8.42	8.32
Southeast Anatolia	5.26	5.28	5.46	5.29	5.38
East Anatolia	5.63	5.20	5.00	4.39	4.15

Source: Gezici, F., Hewings, J.D. (2002).

On the other hand, Marmara is the most developed region, with Istanbul the biggest city of Turkey, being located in this region with its 10 m. population. This region is the heart of the Turkish economy. The Aegean Sea and Central Anatolia regions are also well developed. Generally, the most developed cities are located in the Western regions and along the west and south coast.

Educational services and quality are also very different between the eastern and western parts of the country. The most successful schools of the country are located in the west side of the country. In the high school and university entrance exams, the west part is significantly more successful.

The inequality between regions has led to major migration from the east to west. As a result of the migration, the economic, social, and educational problems have been increasing in big cities. Western migration affects the class size: Large classes are impacting on educational quality.

Turkey has also experienced serious economic problems. The economy has suffered from volatility in the macro economy and chronic inflation for three decades. The annual inflation rate in Turkey has averaged 100 per cent. Since 2002, the government has had an agreement with the

International Monetary Foundation (IMF) and, as a result of new economic policies, inflation fell below 10 per cent in 2004 and macro economical balances were stabilised.

In spite of these positive developments in the Turkish economy, poverty and unemployment are still high and threaten the social peace. According to Turkey State Statistic Institution, the rate of total participation to labour force is 49 per cent and unemployment rate is 9 per cent with agriculture and 13 per cent without agriculture in 2004. The state is dominant in economy and the biggest employer, and privatisation is very slow. In spite of commitments made to the IMF, the governments has not undertaken major privatisation. Moreover, the state economic institutions have been the corruption centres and successive governments have used these institutions to make rich their supporters. A profile of the Turkish economy is presented in Table 2.

Table 2. Key figures in Turkish economy

GDP	\$183 billion
GDP/capita	\$2,700
GDP/sector	Agriculture 13%, Industry 30%, Services 57%
Foreign debts	\$118 billion (2001)
Foreign debts/GDP	64%
Foreign debts/capita	\$1,760
Annual growth in GDP	4.9%
Trade balance	-19%
Annual inflation rate	45%
Official unemployment	10.8%

Source: Encyclopaedia of the Orient (All figures are 2002 estimates) http://i-cias.com/e.o/turkey_2.htm

Turkey has the lowest per capita in GDP among EU states according to purchasing power parities. With average per capita GDP in EU set at is 100, Turkey's index is 28, Germany is 108, Greece is 81, Portugal is 74, Poland is 46, and Romania is 30 (Eurostat, 2003).

The Turkish economy has recovered during last two years (2003-2004), growing 6 to 7 per cent yearly, which is the fastest rate in OECD countries. The government ran the highest primary surplus on record in that country – just over 6 per cent of Gross National Product (GNP). Turkey needs to continue having large primary surpluses to reduce its debt. Although gross public debt of the new member states of EU is about 40 per cent of their Gross Domestic Product (GDP), Turkey's gross debt is double with 80 per cent (www.imf.org).

Economic problems obviously affect educational investment negatively. In 2004, the first time, the budget of Educational Ministry surpassed the budget for National Defence, but it is not enough to solve educational problems. In the OECD countries, the average educational expenditure per student (primary to tertiary) was \$6,361 in 2000. While the expenditure was more than \$8,000 per students in the United States, Austria, Denmark, Norway, and Switzerland, Turkey had the lowest expenditure among OECD countries¹ with only \$1,000 per student. Obviously, Turkey should increase expenditure per student but government funds are not available for this. In 2000, Turkey had 1.4 per cent of private expenditure in all levels on educational institutions. This was very low when compared with other countries. For example, Table 3 shows that it was 18.9 per cent in Germany and 31.8 per cent in the United States (*Education at a Glance*, 2003). The private sector could contribute to educational expenditure if the government established good regulations for the private sector.

These social, cultural, and economical problems certainly affected the education system. Turkey had the lowest enrolment rates at primary and secondary level among the OECD countries of 30

¹ *OECD Countries:* Austria, Australia, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxemburg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.
Non OECD Countries: Argentina, Chile, India, Indonesia, Israel, Jamaica, Paraguay, Thailand, Uruguay.

per cent for 15-19 year olds, while Germany had 89.4 per cent, the United States 77.6 per cent, Zimbabwe 38.3 per cent, and Uruguay 65.9 per cent. For the 5 to 14 year old age groups, Turkey had the lowest enrolment rate with 83.5 per cent among the OECD countries. According to multiple comparisons of mean performance of fourth Grade students on the Progress in Reading Literacy Study (PIRLS), Turkey had the lowest reading rates among OECD countries in 2001 (*Education at a Glance*, 2003).

Table 3. Relative proportions of public and private expenditure on educational institutions for level of education (2000)

Country	Public Sources	Private Sources
Germany	81.1	18.9
United States	68.2	31.8
United Kingdom	85.2	14.8
Korea	59.2	40.8
Greece	93.8	6.2
Turkey	98.6	1.4

Source: *Education at a Glance*, 2003 (OECD Indicators)

PRIVATISATION AS A NEW POLICY TOOL

When modern Turkey emerged after World War I, the country was very poor and did not have enough capital for investment and the Government played an important role in investments to produce and deliver services. Today, although there is a significant private sector, the Turkish government is still the biggest employer in the country and most of the state-run institutions are losing money because of their ineffective management. Governments started to privatise as a requirement of the market economy and liberalisation. Privatisation had many advantages. For example, consider a state-owned and state-run factory which was losing money. If the government sold this factory, it would stop losing money. Secondly, state would be paid money for selling the factory. Thirdly, the private-run factory would start to make a profit and would pay taxes to the state. Finally, the quality of service and customer satisfaction would change. In this sense, privatisation would be a totally useful policy tool for everybody.

As a policy tool, privatisation involves a transfer of services, property and responsibilities from the government to the private sector. Privatisation is a related concept with 'liberalisation' and 'marketisation' in that private organisations exist as alternatives to government services or state allocation systems (Levin, 2001). Privatisation is seen to maximise efficiency and increase service quality while decreasing costs. The main purpose of the private sectors is to make profit. In order to do this, they have to provide customer satisfaction, efficiency, transparency and quality at the lowest cost. In addition, they have to convince customers about the quality of their services.

Governments can create mechanisms for facilitating private participation in schooling. The private sector can be the owner and operator of schools. But the private sector can also operate public schools under contract. Under this approach, government can contract with education colleges, and education unions for more effective management. The private sector can also provide educational materials such as textbooks, and operate buildings, canteen services, and transportation services. This approach is very common in Turkey. In addition, if governments make good regulations involving a core curriculum and transparency of administration, religious groups, diverse cultural groups, and some organisations can build and operate schools. In this way, they can share the educational cost of governments and service the need for diversity in society.

The World Bank and OECD are advocating privatisation policies for developing countries to address their educational problems. Belfield and Levin (2002) explain that "privatisation in education eases the pressure on governments to meet increasing demand and relieves them of excessive cost" (p. 7). Education is a very expensive investment in both developed and

developing countries and government sources alone are not enough to provide quality education for all students. Privatisation eases some of these stresses. The private sector can be involved in educational investment to build and run schools, if they are supported by good regulation.

Private schools can also meet the need of diversity in society. Many ethnic and religious groups seek a different form of education than the one offered in public schools. Public schools have standardised practices and cannot meet specific group or individual needs. In principle, private schools do not need to follow the same calendars, course books or curriculum as do public schools.

Private schools are successful because they are more accountable. In private schools, teachers have no permanent status. If they do not work well enough, teachers can be fired. School owners and principals are also accountable to parents, because parents can withdraw their children if they do not obtain sufficient satisfaction. In Turkish public schools, teachers are accountable to only the principals and inspectors; and principals are accountable to district management; and district managements are accountable to the Educational Ministry; but nobody is accountable to parents and children. On the other hand, in the private sector, parents' and students' satisfaction are very important and all staff are accountable to parents and children.

PRIVATE EDUCATION PRO AND CON

In Turkey, education is undertaken under the supervision and control of the state. According to Article 42 of the Constitution, everyone has the right to receive education. Primary and secondary education is free of charge in public schools.

All private schools in Turkey are subject to Law No. 625 and fall under the administration and inspection of the Ministry of National Education. There is no curriculum freedom for private schools. All curricula are prepared by Educational Ministry and private schools must use them. Curricular innovations in schools are impossible.

In Turkey, most private schools can be classified as profit-making. Profit-making schools are opened as a result of diversification and unmet demands. Generally, they service middle- and upper- class clienteles in urban area. Some private schools have an Islamic character. Religious Islamic schools are not allowed and so religious schools do not say that they have an Islamic character. Customers know which schools are Islamic or secular.

As seen as Table 4, enrolments in private schools are low. The percentage of students in private schools (preschools, primary education and secondary education) is 1.6 per cent.

Beginning in the 1990s, the Turkish government began planning to increase the ratio of private education to 6 per cent in the short-term and to 15 per cent in the long-term. Government allocated land to 12 private schools in 1992 and supported 258 people and institutions from 1991 to 1998 to open private schools (Istanbul Ticaret Odasi, 1999 pp. 17-18). In addition, in summer 2004, the Government waived taxes for private schools for the next three years.

Public sector advocates have opposed the expansion of private sector. According to these advocates, the goal of privatisation was an increase in the role of parents in the financing of education, which could increase inequalities in access to education and break social cohesion (Belfield and Levin, 2002). Education should be for the public good, but private education sectors attached more importance to markets instead of the public good. As a result, any expansion of private education would be at the expense of the social efficiency and equity. In addition, public sector advocates have said that some parts of education was an inevitable responsibility of government and private sector was unwilling to invest in such areas. Another problem was a lack of information for consumers in the education market.

Table 4. Comparing private and non-private education in Turkey (2002-2003 education year)

Education level	Total Number of Students	Total Number of Teachers
All preschools	320,038	18,921
Private preschools	(3%) 9,759	(3,3%) 626
All primary schools	10,331,619	390,275
Private primary schools	(1,5%) 155,868	(3,7%) 14,740
All secondary schools	3,034,959	148,563
Private secondary sch.	(1,8%) 57,365	(4,2%) 6,375
All schools	13,686,616	557,759
All private schools	(1,6%) 222,992	(3,8%) 21,741

Source: Translated and arranged from official web site of Turkish Educational Ministry
<http://www.meb.gov.tr/stats/apk2003/icindekilerSayisalVeriler2003.pdf>

Private sector advocates claimed that government did not need ownership of schools to overcome these problems. Thus, according to the World Bank (1999), these problems could be solved by effective regulation. If the Government established an appropriate mechanism to control quality and ensure transparency and accountability by forcing private schools to disclose information periodically, then the problem could be solved (pp. 6, 16, 17).

Private schools can play an important role in Turkish education and most policy makers support the expansion of the private sector in education. But such expansion requires the combined efforts of the public and private sectors.

HOW SHOULD PRIVATISATION BE SUPPORTED?

According to Tooley (2002), the role of governments should be changed for four main reasons. First, public schools are not managed efficiently and effectively. Second, there are doubts about the accountability of public schools, especially to poor families. Third, a competitive environment increases education quality. Fourth, the private sector can provide important alternative sources of educational funding (p.1). Instead of providing all education services, governments only provide priority services that cannot be purchased in the market. Governments can establish accountability and accreditation systems and turn over some or all schools to contractors. Governments can share the private school expenses to expand the privatisation in education. For example, Government can pay teachers' salaries or the running costs of private schools or Governments can pay the tuition of students. In addition, private educational investments can be tax-free because they are helping the government by investing money. In this way, private schools can replace all or most public schools. Private schools can increase the quality in education and meet diversity the necessary in society through effective regulation.

Regulation, funding, and accountability are the most effective ways to support privatisation. Through regulation, governments can establish a clear legislative framework and specify the roles of the public and private sectors. One of the important issues in regulation centres on student admissions. Schools can choose students based on independent criteria or schools must give preference to families of a certain type or location; admission can be done by lottery. According to the World Bank (1999, p.15), governments need to deal with such issue so that families who have different financial resources and needs can access education. In addition, schools need to meet basic quality standards and contribute social cohesion. Beyond a set of such issues, governments need to provide freedom for private schools to ensure their effective management.

Funding is another policy tool. A funding system determines who pays education costs to the private sector. The general approach is that the person who benefits needs to pay but education provides social benefits in addition to individual benefits. This yields a new question: What proportion of the overall costs of education is best paid for by students and by governments. There is general agreement that public benefits are greatest at earlier stages of education and less

at later stages, so most government funding is best directed at early childhood and primary education and less government funding is directed at secondary and higher education levels (World Bank 1999, pp. 17, 18).

Governments can purchase goods and services for students instead of providing these services. In Turkey, the Government provides the 98.5 per cent of educational services by hiring teachers, building and running schools. In Turkey, private schools are supported by the payment of tuition fees. Governments can share the educational cost of private schools and, by doing so can help privatisation expand.

Vouchers and scholarship are common ways to fund education in the private sector. Governments can pay for students, who attend private school (a voucher system). In the voucher system, governments can determine eligibility of students. Vouchers can be made available for all students or only for low-income students. Alternatively, vouchers can be made available for disadvantaged groups, for example; girls in rural areas, disabled students. Governments can use voucher policy for schools such as religious schools or technology-based schools. Voucher amount is another important issue. Voucher amount² can be equal to per-pupil expenditures in public schools or less or equal to private school fees. School choice is an inevitable feature of private education but it creates a transportation problem. A voucher system can cover all transportation expenditure, some of it, or any part of it. All these policy options associated with voucher systems depend on government policy and country conditions. Levin (2001) explains that a good voucher policy should support four criteria: freedom of choice, productive efficiency, equity, and social cohesion. Vouchers are clearly seen to be a policy tool to expand private education

Accountability is another policy tool that is necessary for transparency. Consumers need information for making effective choices. If the private sector does not provide enough information, governments should do it.

Policy makers can mix these policy tools in an optimal way to meet government objectives in education (World Bank 1999, p.15). Governments need to create national standards, an accreditation system, and transparency and provide freedom to private schools to shape their own curriculum and innovations. Schools need to have a choice in curriculum, teachers, salaries, textbooks, and budget. In these ways, governments can provide competitive schools, and such schools can be monitored by parents and students who receive the services directly. This is a more effective control system because parents can withdraw their children from private schools if they are not satisfied or demand a better service and teachers can be fired if they are not teaching well. Thus, privatisation can give parents more choice and governments can regulate private schools, inspect them for basic standards, and create an accountability framework including a teachers' licence, a core curriculum, and academic standards for students.

Private schools share educational expenses with government by spending money for buildings, and for operational expenses. If these schools are closed government would pay for all these expenditures. If government were to support private schools, the number of the private schools would be increased and government can save money from educational expenditure³. The money can be used for improvement of other public schools and for funding disadvantaged students.

² Ahmet Simsek, an owner a private high school in Istanbul/Turkey, suggests that private schools can service students from public schools if government pays them the same cost as they spend per student (Istanbul Ticaret Odasi 1999, p.54). Unfortunately, Cevdet Cengiz, who is the associate undersecretary in educational ministry, opposes the suggestion. According to him, it is not applicable but he does not explain why (Istanbul Ticaret Odasi, 1999, p.55). A voucher system, which is the most effective supportive system in private schools, is not known or not used in Turkey.

³ Teachers and teacher unions probably oppose the expansion of private schools because of their job security. In the current Turkish education system, once you become a teacher it is almost impossible to be fired. Governments need

PRIVATE SCHOOLING IN ANKARA

Ankara is the capital and the second largest city of Turkey with a population of 3.5 m. in 2003. Ankara is an important commercial and industrial city and serves as the marketing centre for the surrounding agricultural area.

I decided to describe the availability of private schooling in Ankara as a baseline for thinking about the development of private schooling in Turkey. The city can also characterise the country. Thus Ankara is the major population centre with every kind of subculture group; Kurdish, Laz, Alavi, Muslim, poor, rich, educated, and illiterate.

Ankara has 24 districts. The number of schools, students, and teachers in Ankara is presented in Table 5. Cankaya is the largest district and contains the most of the private elementary schools. Kecioren is the second largest district with 13 private elementary schools. Yenimahalle is third largest district and has four private elementary schools. Cubuk, Etimesgut, and Polatli each have two private elementary schools. Mamak, Golbasi, Altindag, Beypazari, and Sincan each have one private elementary school. The remaining districts do not have any private elementary schools.

Table 5. The number of schools, students, and teachers in Ankara (2002-2003 edu. year)

	Preschools	Elementary Education	Secondary Education
The number of schools	818	1,018	328
The number of students	25,002	549,889	182,338
The number of teachers	2,037	23,419	12,067

Source: Translated from official web site of Turkish Educational Ministry
<http://www.meb.gov.tr/stats/apk2003/SayisalVeriler2003.htm>

Population and socio-economic situation can explain the dispersion of private elementary schools in districts of Ankara. As seen in Table 6 and in Figure 1, there is a significant positive relationship between school numbers and the size of the population.

On the other hand, there are some exceptions and it is possible to explain these exceptions by drawing on an understanding of the socio-economic situations (SES) of districts. Table 6 also shows that Ankara has a total of 60 private elementary schools. More than half of the private elementary schools (32) are located in Cankaya. It is not possible to this just by its population. The expected number of private elementary schools is 12 but it has 32 private elementary schools. This can be explained by the socio-economic situation of Cankaya which has the best socio-economic condition among districts of Ankara.

Likewise, Kecioren and Cubuk have more private schools than expected. Both districts are high SES. Kecioren also has more private elementary schools than expected. The reason for this can be explained by its good socio-economic condition. The current Prime Minister is also living there with his family. Cubuk has two private elementary schools which is one more than expected. Some politicians and rich persons were living in Cubuk and their economic conditions were very good according to Turkey conditions.

Also shown in Figure 1, Yenimahalle, Mamak, Altindag, and Sincan, have fewer private elementary schools than expected in terms of their population sizes. All these districts are known as less-developed and their socio economic situation is very low. For example, many residents of Mamak and Altindag are from rural areas around the city and from the east-side of the country. They are poor people. Their living conditions are very hard, in many cases with eight to 10 people live just in one room.

to think about the students' right instead of teachers' rights and provide a consumer-centred education by expanding private education.

Table 6. Population, socio-economic level (SES) and the number of private elementary schools in the districts of Ankara

Districts	Population	Socio-economic Development level	Number of Private Elementary Schools	Expected Number of Schools According to Population Ratio	
1. Cankaya	769,331	N/A	32	12	>
2. Kecioren	672,817	N/A	13	10	>
3. Yenimahalle	553,344	N/A	4	8	<
4. Mamak	430,606	N/A	1	6	<
5. Altindag	407,101	N/A	1	6	<
6. Sincan	289,783	N/A	1	4	<
7. Etimesgut	171,293	N/A	2	2 or 3	=
8. Polatli	116,400	2	2	2	=
9. Cubuk	75,119	4	2	1	>
10. Golbasi	62,602	1	1	1 or 0	=
11. S.Kochisar	59,128	7	0	1 or 0	=
12. Haymana	54,087	17	0	1 or 0	=
13. Beypazari	51,841	6	1	1 or 0	=
14. Elmadag	43,374	3	0	1 or 0	=
15. Nallihan	40,677	10	0	1 or 0	=
16. Bala	39,714	16	0	1 or 0	=
17. Kizilcahamam	33,623	11	0	1 or 0	=
18. Kazan	29,692	5	0	0	=
19. Kalecik	24,738	14	0	0	=
20. Ayas	21,239	9	0	0	=
21. Gudul	20,938	13	0	0	=
22. Akyurt	18,907	12	0	0	=
23. Camlidere	15,339	8	0	0	=
24. Evren	6,167	15	0	0	=
TOTAL	4,007,860		60		

N/A = These are central districts and there is no reported information about their socio-economic conditions.

> More than expected schools according to population ratio;

= Expected schools according to population ratio

< Less than expected schools according to population ratio

Source 1: State Institute of Statistics (SIS) Prime Ministry Republic of Turkey (2000, p.59)

Source 2: Prime Ministry, Government Planning Department. <http://ekutup.dpt.gov.tr/bolgesel/dincerb/ilce/2.pdf>

For all other districts, there is a strong relationship between their population and the number of private elementary schools. Ayas, Bala, Beypazari, Camlidere, Elmadag, Evren, Gudul, Haymana, Kalecik, Kazan, Kizilcahamam, and Nallihan do not have any private elementary schools. This is not surprising because all of these districts have small populations and their socio-economic situation is low. Etimesgut, and Polatli have two private elementary schools and Beypazari and Golbasi have one. In terms of their population, these numbers are expected.

In summary, private elementary schools in Ankara are dispersed according to the socio-economic situations of the districts. Most of these schools are profit-making and service the middle- and upper- class families to meet their quality education and religious needs. Private schools are not accessible for poor families. This causes inequality between rich and poor students in high schools and university entrance examinations, so that disadvantaged groups who are the majority of population, oppose the privatisation. Privatisation can help to solve many educational problems if government regulates it in ways that make private schooling accessible to students at different income levels.

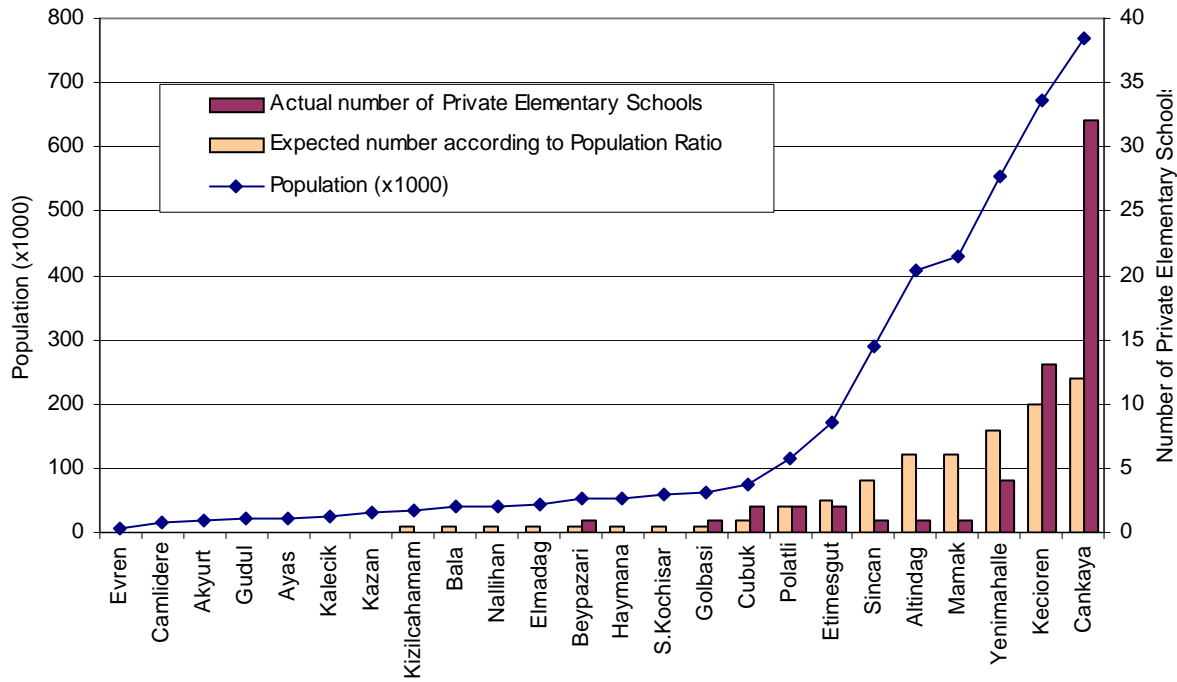


Figure 1. The relationship between the population of districts and the actual number of private elementary schools in Ankara, along with the expected number of schools

CONCLUSIONS

Turkey, as a developing country, faces many problems in education. Large class size, low enrolment rate, girl's education, high illiteracy rate, religious education, textbooks, curriculum and multicultural education are some of the important problems. On the other hand, cultural diversity, economical problems (lack of resources), migration, central management, the rapid growth in school-age population and regional differences make the solutions very difficult.

In principle, privatisation as a policy tool in education can help in the solution of these educational problems in Turkey. Private schools can share responsibility with the Government by investing money in education. In addition, private schools can also meet the needs for diversity in society. Private schools can promote specific religious or philosophical views by using their own teaching methods.

However, in Turkey, the role of private education is very limited. Only 1.6 per cent of total educational expenditure comes from private sources when compare to the OECD mean of 11.6. There is no curriculum freedom for private schools in Turkey. Private schools do not need to follow the same calendars, course books or curriculum like public schools because they also have a mission to meet the necessary diversity and implementation of innovations. The Turkish Government can expand the privatisation in education to solve its educational problems. The Government needs to create an enabling environment, to increase accountability and transparency, and to establish an appropriate mechanism to control quality.

REFERENCES

- Bauman, C.B. (1996). *Governing Education: Public Sector Reform or Privatization*. Needham Heights, MA: Simon and Schuster.
- Belfield, C.R., Levin, H.M. (2002). *Education Privatization: Causes, Consequences, and Planning Implications*. Paris: UNESCO.

- Cassidy, E., Delhaxhe, A., Sacre, A. (2000). *Private Education in the European Union*. Brussels: Eurydice European Unit.
- Dinler, A., Subasi, B. (2003). *Dunyada ve Turkiye’de ozel okullar [Private schools in the world and Turkey]*. Istanbul: Istanbul Ticaret Odasi.
- Encyclopaedia of the Orient, (2002). *Key Figures in Turkish Economy* [Online] http://i-cias.com/e.o/turkey_2.htm [2004, September 7].
- Finkel, A. Sirman, N. (1990). *Turkish state, Turkish society*. New York: Routledge.
- Gallacher, N. (2001). *Governance for quality of education: Conference proceeding*. Budapest: World Bank and Open Society Institute.
- Governor of Ankara (2004). *Problems Concerning Education*. [Online] <http://www.ankara.gov.tr/index.php> [2004, October 10].
- Istanbul Ticaret Odasi (1999). *Ulusal egitimde ozel okullarin yeri ve sorunlari [The Value and Issues of Private Schools in National Education]*. Istanbul: Tanburaci matbaacilik.
- James, T., Levin, H.M. (1988). *Comparing private and public schools*. New York: Falmer.
- Levin, H.M. (2001). *Privatizing Education. Can the Market Deliver Freedom of Choice, Productive Efficiency, Equity and Social Cohesion?* Boulder, CO: Westview.
- Levin, H.M. (2003). *The Marketplace in Education*. New York: Teachers College.
- McEvan, P.J. (2000). *Comparing the Effectiveness of Public and Private Schools: A Review of Evidence and Interpretations*. New York: Teachers College.
- Metz, H.C. (1995). *Turkey a Country Study*. Washington, D.C: U.S. Government printing office.
- OECD. (1989). *Reviews of National Policies for Education: Turkey*. Paris: OECD.
- OECD. (2003). *Education at a Glance: OECD Indicators*. Paris: OECD.
- OECD. (2002). *Education Policy Analysis*. Paris: OECD.
- OECD. (2002). *Turkey: Crucial Support for Economic Recovery*. Paris: OECD.
- Ozgen, B. (1996). *Turkiye’de parali egitim. [Education with Money in Turkey]* Izmir: ECEV.
- Prime Ministry Government Planning Department (2004). *Ilcelerin sosyo-ekonomik durumlarina gore siralanmasi*. [Online] <http://ekutup.dpt.gov.tr/bolgesel/dincerb/ilce/2.pdf> [2004, October 8].
- Randall, E. V. (1994). *Private Schools and Public Power: A Case of Pluralism*. New York: Teachers College.
- Roberson, B.A. (1998). *The Middle East and Europe: The Power Deficit*. New York: Routledge.
- State Institute of Statistics Prime Ministry Republic of Turkey. (2001). *Census of Population*. Ankara-Turkey: State Institute of Statistics.
- Tooley, J. (1999). *The Global Education Industry: Lesson from Private Education in Developing Countries*. Westminster, London: The institute of economic affairs.
- Tooley, J. (2003). *Government Failure: E.G. West on Education*. Westminster, London: The institute of economic affairs.
- Turkish Educational Ministry (2003). *SayisalVeriler*. [Online] <http://www.meb.gov.tr/stats/apk2003/SayisalVeriler2003.htm> [2004, September 13].
- Turkish Educational Ministry, Ankara District Management (2004). *Ozel ogretim*. [Online] <http://ankara.meb.gov.tr> [October, 16].
- World Bank (1999). *EdInvest: Private Education*. [Online] <https://www.worldbank.org/edinvest> [2004, November 5].

Does increasing communication through visual learning environments enhance student perceptions of lecturers?

Lara Frumkin

Middlesex University, London frumkinl@mdx.ac.uk

The current study was conducted in an effort to examine whether increased levels of communication using visual learning environments (VLEs) alters student perceptions of lecturers. Eighty-six MSc students in Computing Science participated by using She and Fisher's (2002) Teacher Communication Behavior Questionnaire (TCBQ). In addition to using the questionnaire, data from the electronic class site were used to make assessments about the quality and quantity of communication. Two types of classrooms were evaluated: a) a control condition in which the lecturer did not alter any communication aspect of the module, and b) the experimental condition in which the lecturer posted weekly discussion topics. Significant differences were found by cultural background and gender of the students. The bulletin board postings in the experimental condition were more heavily content-based than the control condition ones. The consistency in discussion topic of the experimental condition postings, both bulletin board and email, were more fluid than in the control condition.

E-learning, visual learning environment, communication, culture, student perceptions, gender

INTRODUCTION

Online education is becoming increasingly popular in academia (Collis, 1996; Dutton, Dutton, and Perry, 2002; Hiltz and Wellman, 1997; Schweizer, Paechter, and Weidenmann, 2001; Spitzer, 1998; Stadtlander, 1998; Webster and Hackley, 1997). In academia in the United States alone, of the 5655 accredited postsecondary education institutions, 1979 offer a form of distance delivery (Council of Higher Education Accreditation, 2002). The vast literature in the field indicates that the distance education programs are being studied extensively (for example, Buerck, Malmstrom, and Peppers, 2002; Dutton, Dutton, and Perry, 1999; Frumkin, Mimirinis, Dimitrova, and Murphy, 2004; Russell, 1999).

It is important to assess how teachers' behaviours are perceived by students since 63 per cent of what happens in a classroom may be explained by the student's perception of the teacher's influence, a factor possibly based on teacher behaviour (van Tartwijk, 1993). Past work has found that perceptions of behaviours can influence affective learning (Anderson, 1979), cognitive learning (Gorham, 1988; McCroskey, Sallinen, Fayer, Richmond, and Barraclough, 1996; Richmond, Gorham, and McCroskey, 1987; Richmond, McCroskey, Kearney, and Plax, 1987), the effectiveness of the teacher (Anderson, 1979; Cheng and Tsui, 1996), interactions between the student and teacher (Andersen, Norton, and Nussbaum, 1981; Choi, 2002; Moller, 1998), and general student performance (Matsumoto, Garside, and Roberts, 1991; Picciano, 2002).

Student motivation is an important aspect of the learning process (Hall, 1966). Christophel (1990) defines student motivation as the process of 'how' students are taught, rather than 'what' it is that they are actually taught. Vansteenkiste, Simons, Lens, Soenens, Matos, and Lacante (2004)

argued that in the classroom, increasing the quantity of motivation might positively alter the quality of the learning experience regardless of whether the motivation was intrinsic or extrinsic. Ryan and Deci (2000) argued that intrinsic tasks could be viewed positively even if a student was not interested in the task for its own sake. Rather, it could be the intrinsic desire for some future that the task led to one that facilitated the accomplishment of the task. In fact, it has been claimed that the value one placed on a task predicted academic achievement (Eccles and Wigfield, 2002). The more readily a task (for example, learning a computer programming language) led to a future goal (for example being offered a high paying computer program job), the more motivated a student will likely be to complete the task (Eccles and Wigfield, 2002; Lens, Simons, and Dewitte, 2002; Miller, DeBacker, and Greene, 1999). However, lecturers might be able to extrinsically motivate a student to do well (for example, success on coursework results in high grades, possibility of internships).

The study described in this paper uses a framework of extrinsic motivation. It evaluates how distance students perceive teacher communication behaviours. The current study assesses the relationship between perceptions of communication behaviour and module outcome (that is, how well the student does in the module) among other variables. It is hypothesised that positive perceptions of teacher communication can provide extrinsic motivation for the students such that students who perceive their teachers as more communicative are more likely to do better in the module.

Distance Education

Student academic success is equivalent when measured by final grades in distance-based and lecture-based courses (for example, Dutton, et al, 1999; Russell, 1999). Buerck, et al (2002) report that, specifically for computer science students, those enrolled in online courses performed as well as their traditional program counterparts.

The use of computer-mediated technology might expect to enhance communication (Schweizer, et al, 2001). Schweizer, et al (2001) reported that students acknowledge disadvantages of online education such as missed opportunities in communication, anonymity, and high demand on resources. Interaction was a critical element in the learning process (Moore, 1993; Offir, 2000). Interaction between students and teachers might be even more important in the online learning environment (Gresh and Mrozowski, 2000). While time consuming, research has shown that students wanted to be able to access lecturers in a virtual learning environment (VLE) (Sanders and Morrison-Shetlar, 2001).

Online lecturers must move from the instructor role to the facilitator role (Lin and Hseih, 2001). According to Gates (2000), lecturers need to be able to increase their levels of interaction with online students by using effective pedagogical tools and incorporating innovative design features. The use of forced interaction and discussion on module relevant topics might assist students to embrace the material and making them feel as if they 'belong to the classroom' even if they were in distant locations.

Communication Behaviours

Communication in the classroom is comprised of communication with the instructor and communication, as a separate category, with other students (Anderson and Garrison, 1998). Communication with the instructor allows the student to ask questions but perhaps more importantly, to develop a working relationship on which to base assignments and grading. It also kept the student feeling connected to the academic institution and provided a feeling of proximity to an expert in the field (Miller, Preston, Elbert, and Lindner, 2003). Richmond (1990) believed that there was a link between the way a teacher communicated and the way the student learnt. She further argued that the connection between motivation and learning are critical; that is, those who

were motivated learn more and those who learn become increasingly motivated. Motivation might be triggered by the communication style of a teacher (for example, extrinsically created motivation).

Communication in face to face modules was comprised of spoken verbal information and nonverbal personal or social cues (Schweizer, et al, 2001). Verbal messages provided cognitive aspects of communication, while nonverbal messages appeared to be the ones responsible for affective types of communication (McCroskey, et al, 1996). It has been claimed that nonverbal behaviours were relevant in educational environments because they are relied upon for true sentiment or emotions that are hidden when a verbal-only message is provided (Schweizer, et al, 2001). Eye contact and smiling are positively related to cognitive learning (McCroskey, et al, 1996); Teachers' active participation in school activities assisted students in exercising skills and talents (Vansteenkiste, et al, 2004). However, in an online environment nonverbal communication might not be particularly useful.

Verbal behaviours were useful in such a setting. Samples in the Netherlands, the United States, and Australia revealed that friendly, helpful, and understanding teacher behaviour was connected to higher cognitive outcomes and positive student attitudes (Fisher, Henderson, and Fraser, 1995; Fisher and Rickards, 1997; Wubbels and Levy, 1993).

Challenging communication behaviours such as teacher questioning and reaction to student answers promoted relevance of a given topic, encouraged ownership of module material, assisted students in their interpretations of new module content, and connected recently learned information to information students already have (Deal and Sterling, 1997; Good and Brophy, 1974; Walberg, 1984). She (1998a, 2000, 2001) found that teacher questioning and verbal reinforcement following high performance by students were positive facets of teacher behavior (that is, requiring students to collate new information with already-existing information, encourage ownership of educational material, and assist students in analysing the new content). Furthermore, Comadena, Semlak and Escott (1990) found that among adult learners, a dominant teacher style was a predictor of the teacher's effectiveness.

Research has shown that controlling behaviour on the part of teachers increased cognitive gains among students (Wubbels and Levy, 1993; Fisher, et al., 1995; Fisher and Rickards, 1997). A study looking at Dutch students found that girls perceived their teachers as being more dominant (that is, controlling) than did boys (Levy, Wubbels, and Brekelmans, 1992).

A Taiwanese study revealed that student achievement was increased when students felt that their teacher exhibited behaviours such as encouragement and praise (She and Fisher, 2002). Motivation which might be provided by teacher encouragement, or praise of the student, enhanced interest and involvement in class as well as students reportedly looking forward to attending lectures (Frymier, 1994).

Gender

Past work has found some differences in online student behaviour based on gender. Males preferred to work independently to develop more class postings and were more likely to ask lecturers for assistance. The females preferred more classroom interaction, use language that was more complimentary when responding to other learners, and were more likely to ask peers for help (Trego, 2004). Australian male students preferred an individualised classroom when compared with their female counterparts (Hansford and Hattie, 1989). Taiwanese girls more frequently than boys reported their teachers as being understanding and friendly (She and Fisher, 2002). Several studies indicated that girls perceived their learning environments more positively than did boys, regardless of cultural background (Fraser, et al, 1995; Fisher, et al, 1997; Rawnsley and Fisher, 1997). Since females perceived their learning environments more positively and

enjoyed greater levels of interaction, it was logical to assume that females who felt good about an instructor would give him or her higher ratings on his or her communication skills.

Culture

The cultural background of the student might affect how the particular student perceived his or her teacher's behaviour (Mehrabian, 1969; Powell and Harville, 1990). Perceptions might be influenced by a number of behaviours that differed across cultures (McCroskey, et al., 1996). Cultures needed not to be defined by function of being in different parts of the world; rather, cultures should be considered as sociopsychological entities in and of themselves (Lee, Matsumoto, Kobayashi, Krupp, Maniatis, and Robert, 1992). For example, an individual who considers herself Egyptian, even if she was living in London, might have more of an Egyptian cultural personality than an English one.

Teacher behaviours were not related to cognitive learning among United States African-American, Asian-American, Hispanic, and non-Hispanic Caucasian groups of students (Sanders and Wiseman, 1990). In Chinese society the teacher-student relationship might be compared to the father-son one (Pratt, Kelly, and Wong, 1998). The teacher role commanded a certain level of respect from the student. In return, the teacher should exercise authority over the material which he or she was teaching. Pratt, et al (1998) further argued that in Western society teachers might compromise their position as an authority figure in an effort to be better liked. This did not provide a clear picture of what to expect with regard to perceptions of communication patterns based on culture. Nonetheless, it might be assumed, based on Pratt's (1998) work, that Chinese students would be less likely to provide ratings of their lecturers because of the authority position which they held. That is, it might be irrelevant to a Chinese student what he or she thought about a teacher. The teacher was the expert and therefore respected, regardless of communication behaviour.

Current Study

It is expected that online educational communication depended on a number of social and personal variables. The current study was designed to manipulate experimentally an already developed module to investigate whether increased levels of communication using VLEs altered student perceptions of the lecturers. Measurement of whether perceptions of the lecturers' communication behaviours tangibly influenced module outcome scores. The effects of gender and cultural background were also assessed. The framework for this study was that extrinsic motivation was provided by the teacher, both online and in the classroom. As the motivation to succeed externally was provided by the teacher, through enhanced communication behaviour online, the students learnt to incorporate it internally such that they became intrinsically motivated. This led to students having higher levels of academic achievement (that is, module outcome scores would be higher), learning would be enhanced and students would report more positive perceptions of their teachers.

METHOD

Participants

Eighty-six MSc students in Computing Science, 53 Asian and 33 Caucasian, from two modules participated in the study. Only eight of the participants had lived in the United Kingdom for five or more years, while 66 had moved to the United Kingdom in the past year. The remaining 12 students have lived in the United Kingdom between one and five years.

Materials

She and Fisher's (2002) *Teacher Communication Behavior Questionnaire* (TCBQ) was used to assess student perception of variations in communication style with lecturers. The questionnaire had 40 questions, comprised of five scales, each representing a type of communication. They were (a) challenging, (b) encouragement and praise, (c) nonverbal support, (d) understanding and friendly, and (e) controlling communication. It was a Likert style questionnaire and questions were answered by circling 'almost never' (1), 'seldom' (2), 'sometimes' (3), 'often' (4) or 'almost always' (5). Cronbach alpha coefficient was calculated to measure reliability on the five scales. A range of 0.86 to 0.93 on samples in Taiwan and Australia was found (She and Fisher, 2000). Validity, measured by principal components analysis, showed that the 40-item questionnaire was structured based mainly on the factor analysis and in small part on the interviews conducted with the students. Discriminant validity measures ranged from 0.06 to 0.45. These were small enough correlations between the scales to be satisfactory (She and Fisher, 2000).

In addition to using the questionnaire, data from the electronic class site was collected. Following the end of the semester, and after final grades had been posted, communication from the class bulletin board site and email correspondence was downloaded.

Two types of classrooms were evaluated. In the first, a control condition, the lecturer (Lecturer A) did not alter any communication aspect of the module. In the second, the experimental condition, the lecturer (Lecturer B) posted weekly discussion topics. Students were instructed to engage with each other and the lecturer on the discussion topics. The correspondence was looked at both for quantity of contact as well as quality of discussion (for example, questions about coursework or due dates to more substantive questions regarding module content). To ensure that differences were not lecturer-specific only, communication only data (no questionnaires) were obtained from Lecturer B during the enhanced communication semester as well as the semester prior to the modification (the second control condition).

Procedure

Questionnaires were distributed to two groups of students. The first completed the questionnaire in Spring 2004 and the second group in Fall 2004. The principal researcher attended the lecture sessions, in both cases with an associate, to disseminate and collect the questionnaires, and to respond to questions about the research.

Hypotheses

The study models earlier work comparing communication patterns at the secondary school level in Australia and Taiwan (She and Fisher, 2002; She and Fisher, 2000). A framework to investigate student perceptions of communication at the university level was used. It modelled an earlier study with similar students although this time using an experimental design (Frumkin and Murphy, manuscript submitted for publication). Several hypotheses were developed.

- Hypothesis 1: There is a positive relationship between student perceptions of the teacher communication patterns and module outcome; the larger the degree to which a student believes the teacher interacts with the student, the higher the module outcome is for that particular student.
- Hypothesis 2: There are cultural differences with the Asian students reporting less overt patterns of all communication behaviours than the Caucasians due to their respectful nature towards lecturers.
- Hypothesis 3: Female students report greater levels of all communication patterns from their lecturers than male students.

Hypothesis 4: The module with enhanced communication results in greater content, more substantive correspondence and postings and better linked discussion threads than will the module without the enhanced communication.

RESULTS

There were no significant effects found for Hypothesis 1. There was a significant effect for Hypothesis 2. Roy's Largest Root ($F = 3.79$, $p < 0.00$) demonstrates significant differences in perceptions of tutors by cultural background of the students. Additional univariate analyses were run with the cultural background variables. There were significant differences by culture and challenging behaviour ($F = 2.18$, $p < 0.04$), and culture and controlling behaviour ($F = 4.94$, $p < 0.03$) and a significant difference on a non-predicted interaction of culture by gender ($F = 2.46$, $p < 0.05$). Follow-up t-tests revealed findings in the predicted direction for challenging behaviour ($t = -5.91$, $p < 0.01$) such that Caucasian students reporting more challenging behaviour from their lecturers than did the Asian students. The same held true for controlling behaviour ($t = -3.58$, $p < 0.001$). A t-test for culture and gender ($t = -15.86$, $p < 0.00$) showed that the Caucasian females were more willing to report on perceptions of their teachers than were Asian females, Caucasian males or Asian males.

A multiple regression analysis was run to examine the significance of Hypothesis 3. A significant difference was found on gender for encouragement and praise behaviour ($F = 3.51$, $p < 0.04$). A follow-up t-test indicated that, as predicted, females reported higher levels of encouragement and praise than did male students ($t = 2.44$, $p < 0.02$). There was an unexpected interaction effect of gender by tutor on controlling behaviour. Females rated lecturer A as more controlling than lecturer B. While it was predicted that females would in general rate the lecturers as more controlling, there was a distinct difference in females' ratings of the two lecturers ($F = 3.54$, $p < 0.04$).

Hypothesis 4 was partially supported. The bulletin board postings on Lecturer B's enhanced communication module (experimental condition) were more heavily content-based than the postings in the other two modules (Lecturer A and Lecturer B's non-enhanced communication module). This supported the hypothesis. However, the number of postings, irrespective of content, was higher in Lecturer B's non-enhanced communication module (35 postings) than in the enhanced communication condition (34 postings). While this was an insignificant difference, it was noteworthy that the enhanced communication module did not yield a higher number of postings than the non-enhanced condition. Lecturer A's module had far fewer postings, only 15. A comparison of the content quality of the postings revealed that the only increase in communication for the experimental condition was by the students in their emails (see Table 1). The lecturer posted more content-based emails in the control condition. The students posted more content-based bulletin board messages in the control condition. The consistency in discussion topic of the experimental condition postings, both bulletin board and email, were more fluid than in the control condition. That is to say, the control conditions had more disjointed content-based postings while the experimental postings followed a discussion type of flow, confirming an aspect of Hypothesis 5.

DISCUSSION

Hypothesis 1 posited that students would both perceive the lecturer in a communicative way (high on all communication variables) and receive higher outcome scores. This finding was not recorded. It is unclear why this is the case. It is possible that nonverbal communication is heavily relied upon for perceptions of lecturer behaviour (McCroskey, et al., 1996; Philippot, Feldman, and McGee, 1992; Schweizer, et al., 2001). Nonverbal communication is perhaps near impossible to tease out, if it even exists, in the online environment. It might not be feasible, therefore, to

examine whether lecturer behaviour is manifested as nonverbal communication and influences student's perceptions and their module outcome.

Table 1. Qualitative patterns of communication

	Lecturer A (Control)		Lecturer B (Second Control)		Lecturer B (Experimental)	
	Lecturer	Student	Lecturer	Student	Lecturer	Student
Email (Administrative)	2	2	0	1	0	1 (1)**
Bulletin Board (Administrative)	1	12	0	6 (2)*	0	9
Email (Content)	0	0	8	2	0	8 (1)**
Bulletin Board (Content)	1	1	2	16 (2)*	2	12

* (2) Indicates that two students posted messages containing content and administrative material.

** (1) Indicates that one student posted a message containing content and administrative material.

An alternate explanation for the lack of findings is that perception of communication behaviour is not a primary contributor to module outcome scores. It is possible that motivation, or lack thereof, on the part of the student is a far more significant contributor to module outcome score.

There is partial support for Hypothesis 2. The significant multivariate analysis indicated that a relationship between culture and perceptions of communication behaviour exists. What is interesting to note is that the relationship is not significant for all five communication behaviours. Namely, friendly-understanding, encouragement-praise and nonverbal are not related to culture. Feng (1994) has argued that Asian students were not expected to rate the lecturers dramatically in any category, based on their respectful style. However, the Caucasian students were expected to rate the lecturers more dramatically and this did not surface for the three communication behaviours mentioned above. Nonverbal behaviour might not have been significant for the reasons mentioned above. Consequently, there is a need to explore further cultural differences between the Asian and Caucasian students, especially with respect to the two non significant variables. It is possible that the Caucasian students expected their lecturers at the postgraduate level to be more of a friend or peer than a lecturer. Thus, the Caucasian students rate their lecturers as challenging and controlling. The non-predicted finding fits with other hypothesis-based work. Both females and Caucasians can be expected to rate lecturers higher on communication behaviours, which surfaces in this interaction.

The fact that there is one significant finding for Hypothesis 3 warrants further investigation. It is possible that females are less expectant of their lecturers, since as Trego (2004) notes they rely more on peers for help. Thus it is possible that in this study the students were unaware that the lecturers would be as encouraging as they in fact were. At the same time though, if this were the case, one would expect understanding and friendly behaviour also to have surfaced as being significant for the females. An explanation of this finding requires further research.

The fact that the females rated Lecturer A as more controlling than Lecturer B, but not both lecturers as more controlling than did males, is interesting. Lecturer B is female while Lecturer A is male. Perhaps the females felt that Lecturer B was easier to approach or acted in a less hierarchical manner so that they rated her as less controlling. Alternately Lecturer B, who used the VLE with greater frequency, had changed better from the instructor to facilitator role, resulted in lower controlling ratings (that is, traditional lecturers would have been more in charge of the classroom while VLE instructors were moderators) (Lin and Hsieh, 2001). This finding required further investigation on patterns of communication based on gender.

Hypothesis 5 provided information of a different sort. The students were expected to respond eagerly to the experimental condition by posting more content-based (lecture appropriate) messages. It was also anticipated that students would continue on a consistent discussion path initiated by the lecturer. There was not a greater number of postings for the experimental versus the second control condition but the content of the postings in the experimental condition was better linked from posting to posting. The students did not initiate discussions of their own, but

they did respond to the lecturer's content-based postings. While this is a positive outcome, it is unclear whether the preparation and follow-up work required by the lecturer to post weekly discussions is worthwhile. If simply comparing Lecturer A to Lecturer B's experimental condition, leads to the conclusion that the experiment was a success as there were many more postings. Yet, with the inclusion of the control condition from Lecturer B, it appears that the difference is a lecturer difference more so than an experimental one. Recommendations to Lecturer A, and other lecturers who have low participation in online discussions, may be to impose a structure similar to the experimental condition.

CONCLUSIONS

Taken together, this research provides insight into VLE courses. While they do not appear to be detrimental to the student's performance, altering the communication design of the module does not seem to enhance final module grade or student perceptions of the lecturer. All three conditions yield students with similar academic success rates.

Perceptions of lecturer communication behaviour is related to various factors, such as gender and culture. The findings on culture and gender yield interesting results. What surfaces as the most interesting result is the lack of consistent findings across communication behaviours. Research needs to be conducted to determine whether students do not pay much attention to the lecturer's behaviours and whether this varies based on the type of course (undergraduate, postgraduate). Research also needs to be carried out to assess how or if other aspects of enhanced communication, besides posting weekly discussion topics, appeals more to students and increases academic outcome. Finally, ways to increase motivation through enhanced communications need to be investigated.

REFERENCES

- Anderson, J. (1979). Teacher immediacy as a predictor of teaching effectiveness. In D. Minmo, (Ed.), *Communication Yearbook 3*, (pp. 543-559). Beverly Hills, C.A.: Sage.
- Anderson, J., Norton, R., and Nussbaum, J. (1981). Three investigations between perceived teacher communication behaviours and student learning. *Communication Education*, 30, 377-393.
- Anderson, T. and Garrison, D.R. (1998). Distance education for the research university. *Newsletter of the Albert Distance Education and Training Association*, 8, 12-13.
- Buerck, J.P., Malmstrom, T., and Peppers, E. (2002). Learning environments and learning styles: Non-traditional student enrolment and success in an internet-based versus a lecture-based computer science course. *Learning Environments Research*, 6, 137-155.
- Cheng, Y.C. and Tsui, K.T. (1996). Total teacher effectiveness: new conception and improvement. *International Journal of Educational Management*, 10, 7-17.
- Christophel, D. M. (1990). The relationships among teach immediacy behaviours, student motivation, and learning. *Communication Education*, 37, 323-340.
- Collis, B. (1996). *Telelearning in a Digital World. The Future of Distance Learning*. London: International Computer Press.
- Comadena, M. E., Semlak, W. D., and Escott, M.D. (1990). Communication style and teacher effectiveness: A comparative study of the perceptions of adult learners and traditional undergraduate students. Paper presented at the annual meeting of the Speech Communication Association, Chicago, IL.
- Council for Higher Education Accreditation. (2002). *Accreditation and Assuring Quality in Distance Learning*. Council for Higher Education Accreditation. (Monographs Series, No. 1). Council for Higher Education Accreditation. [Online] http://www.chea.org/pdf/mono_1_accred_distance_02.pdf?pubID=246 [March 1, 2006].
- Deal, D. and Sterling, D. (1997). Kids ask the best questions. *Educational Leadership*, 54, 61-63.

- Dutton, J., Dutton, M., and Perry, J. (1999). Do online students perform as well as lecture students? *Journal of Engineering Education*, 90, 131-136.
- Dutton, J., Dutton, M., and Perry, J. (2002). How do online students differ from lecture students? *Journal of Asynchronous Learning Networks*, 6.
- Eccles, J. S. and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132.
- Feng, J. (1994). Asian-American children: What teachers should know. [Online] <http://www.enc.org/features/focus/archive/multi/document.shtm?input=ACQ-111356-1356,EDO-PS-94-4> [March 1, 2006].
- Fisher, D. and Rickards, T. (1997). A way of assessing teacher-student interpersonal relationships in science classes. Paper presented at the *National Science Teachers Association Annual National Convention*, New Orleans, LA.
- Fisher, D., Henderson, D. and Fraser, B. (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education*, 25, 125-133.
- Frumkin, L.A., Mimirinis, M., Dimitrova, M. and Murphy M. (2004). From e-Learning to b-Learning: How Students Use e-Learning Material in a Blended Learning Environment. In the proceedings of the *E-learn Conference*, Washington D.C.
- Frymier, A. B. (1994). A model of immediacy in the classroom. *Communication Quarterly*, 42, 133-144.
- Gates, G.S. (2000). Teaching-related stress: The emotional management of faculty. *The Review of Higher Education*, 23, 469-490.
- Good, T. and Brophy, J. (1974). *Teacher-Student Relationships: Causes and Consequences*. New York: Holt.
- Gorham, J. (1988). The relationship between verbal teacher immediacy behaviours and students learning. *Communication Education*, 37, 40-53.
- Gresh, K. S. and Mrozowski, S. (2000). Faculty/student interaction at a distance: Seeking balance. *EDUCAUSE*, Nashville, TN.
- Hall, E. (1966). *The Hidden Dimension*. Garden City, New York: Doubleday.
- Hiltz, S. R. and Wellman, B. (1997). Asynchronous learning networks as a virtual classroom. *Communications of the ACM*, 40, 44-49.
- Lee, M., Matsumoto, D., Kobayashi, M., Krupp, D., Maniatis, E., and Roberts, W. (1992). Cross-cultural influences on nonverbal behavior. In R. S. Feldman (Ed.), *Applications of Nonverbal Behavioral Theories and Research* (pp. 239-261). New York: Erlbaum.
- Lens, W., Simons, J. and Dewitte, S. (2002). From duty to desire: The role of students' future time perspective and instrumentality perceptions for study motivation and self-regulation. In F. Pajares and T. Urda (Eds.), *Academic Motivation of Adolescents* (pp. 221-245). Greenwich, CT: Information Age.
- Lin, B. and Hsieh, C. (2001). Web-based teaching and learner control: A research review. *Computers and Education* 37, 377-386.
- Matsumoto, D., Garside, M., and Roberts, W. (1991). Predicting teacher effectiveness: The contributions of teacher performance and emotions. Manuscript submitted for publication in Babad, E. (1992). Teacher expectancies and nonverbal behavior. In R.S. Feldman (Ed.), *Applications of Nonverbal Behavioral Theories and Research* (pp. 167-190). Hillsdale, NJ: Erlbaum.
- McCroskey, J. C., Sallinen, A., Fayer, J. M., Richmond, V. P., and Barraclough, R. A. (1996). Nonverbal immediacy and cognitive learning: A cross-cultural investigation. *Communication Education*, 45, 200-211.
- Mehrabian, A. (1969). Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, 71, 359-372.
- Miller, R. B., DeBacker, T. K., and Greene, B. A. (1999). Perceived instrumentality and academics: The links to tasks valuing. *Journal of Instructional Psychology*, 26, 250-260.

- Miller, R. Preston, A., Elbert, C., and Lindner, J. (2003). Lessons from afar: Concerns of distance students, Presented at the 10th Annual Distance Education Conference, Austin, TX.
- Moller, L (1998). Designing communities of learners for asynchronous distance education, *Educational Technology Research and Development*, 46, 115–22.
- Moore, M.G. (1993). Three types of interaction. In K. Harry, M. John, and D. Keegan (Eds.), *Distance Education: New Perspectives* (pp. 19-24). New York: Routledge.
- Offir, B. (2000). Map for decision making in operating distance learning systems: Research results. *Education Media International*, 37, 9-15.
- Philippot, P., Feldman, R.S., and McGee, G. (1992). Nonverbal behavioral skills in an educational context: Typical and atypical populations. In R.S. Feldman (Ed.), *Applications Of Nonverbal Behavioral Theories and Research* (pp. 191-213). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Picciano, A.G. (2002). Beyond student perceptions: Issues of interaction, presence, and performance in an online environment. *Journal of Asynchronous Learning Networks*, 6, 21-40.
- Powell, R.G. and Harville, B. (1990). The effects of teacher immediacy and clarity on instructional outcomes: An intercultural assessment. *Communication Education*, 39, 369-379.
- Pratt, D.D., Kelly, M., and Wong, W. (1998). The social construction of Chinese models of teaching. Presented at the 1998 Adult Education Research Conference, San Antonio, TX.
- Richmond, V. P. (1990). Communication in the classroom: Power and motivation. *Communication Education*, 39, 181-195.
- Richmond, V. P., Gorham, J. S., and McCroskey, J. C. (1987). The relationship between selected immediacy behaviours and cognitive learning. In M. McLaughlin (Ed.), *Communication Yearbook 10*, (pp. 574-590). Beverly Hills, CA: Sage.
- Richmond, V. P., McCroskey, J. C., Kearney, P., and Plax, T. G. (1987). Power in the classroom VII: Linking behavior alternation techniques to cognitive learning. *Communication Education*, 36, 1-12.
- Russell, T. L. (1999). *The No Significant Difference Phenomenon*. Raleigh, NC: North Carolina State University.
- Ryan, R. M. and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist*, 55, 68-78.
- Sanders, D.W. and Morrison-Shetlar, A. I. (2001). Student attitudes toward web-enhanced instruction in an introductory biology course. *Journal of Research on Computing in Education*, 33, 251-262.
- Schweizer, K., Paechter, M., and Weidenmann, B. (2001). A field study on distance education and communication: Experiences of a virtual tutor. *Journal of Computer-Mediated Communication*, 6. [Online] <http://www.ascusc.org/jcmc/vol6/issue2/schweizer.html> [March 1, 2006].
- She, H. C. (1998a). Interaction between different gender students and their teacher in junior high school biology classes. Proceedings of the *National Science Council, part D: Mathematics, Science, and Technology Education*, 8, 16-21.
- She, H. C. (2000). The interplay of a biology teacher's beliefs, teaching practices and gender-based student-teacher classroom interaction. *Educational Research*, 42, 28-39.
- She, H.C. (2001). Different gender students' participation in the high- and low- achieving middle school questioning-oriented biology classrooms in Taiwan. *Research in Science and Technological Education*, 19, 147-158.
- She, H.C. and Fisher, D. (2000). The development of a questionnaire to describe science teacher communication behaviour in Taiwan and Australia. *Science Education*, 84, 706-726.

- She, H.C. and Fisher, D. (2002). Teacher communication behaviour and its association with students' cognitive and attitudinal outcomes in science in Taiwan. *Journal of Research in Science Teaching*, 39, 63-78.
- Spitzer, D. R. (1998). Rediscovering the social context of distance learning. *Educational Technology*, 38, 52-56.
- Stadtlander, L. M. (1998). Virtual instruction: Teaching an online graduate seminar. *Teaching of Psychology*, 25, 146-148.
- Trego, C. D. (2004). Gender differences in communication patterns and learning styles in asynchronous distance education. (Doctoral dissertation, Walden University, 2004). *Dissertation Abstracts International* 64, 2387A.
- van Tartwijk, J. (1993). *Sketches of Teacher Behavior*. Utrecht: W.C.C.
- Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., Matos, L., and Lacante, M. (2004). Less is sometimes more: Goal content matters. *Journal of Educational Psychology*, 96, 755-764.
- Walberg, H. J. (1984). Improving the productivity of American schools. *Educational Leadership*, 41, 19-27.
- Webster, J. and Hackley, P. (1997). Teaching effectiveness in technology-mediated distance learning. *Academy of Management Journal*, 40, 1282-1309.
- Wubbels, T. and Levy, J. (1993). *Do you know what you look like? Interpersonal Relationships in Education*. London: Falmer Press.

Acknowledgement: The author wishes to thank Geetha Abeyasinghe for access to her students and feedback on the manuscript. I would also like to acknowledge Mike Mimirimis and Paul Smith for their assistance with data collection and entry.

Individual and flexible: Working conditions in the practice of Swedish distance-based teacher education

J. Ola Lindberg

Mid Sweden University, Sweden ola.lindberg@miun.se

Anders D. Olofsson

Umeå University, Sweden anders.d.olofsson@fceduc.umu.se

This article reports on the working conditions within Swedish ICT-supported distance-based teacher education. Data collected from teacher trainees are analysed and discussed in relation to Swedish governmental policies concerning teacher education and distance education and theories emphasising the importance of social aspects of education. The findings indicate working conditions that are mainly controlled by the teacher education program, and that teacher trainees to a high degree are fostered into individualism. Exceptions are in group work, which on the other hand seems to be given only minor attention in teacher education. This raises questions related to the intentions of teacher education. First, there are questions concerning issues of flexibility and choice, more precisely about what aspects are flexible or not. Second, there are questions concerning possibilities of teacher education providing an education that enables teacher trainees to develop the competencies needed to be able to teach.

Teacher education, working conditions, flexible learning, online learning community, Sweden

INTRODUCTION

In the current debate concerning teacher education in Sweden, two issues have been in focus. One is the expected shortage of educated teachers. In trying to solve this dilemma, teacher education institutions have in recent years increased the number of programs offered on a distance basis. Relying on distance education, intentions have been to attract future teachers that otherwise would not attend a teacher education programs due to, for example, geographical location and family situation. Another issue is the teacher education reform of 1999, which has just recently been evaluated by the National Agency for Higher Education. In the evaluation report teacher trainees are claimed to have been left without much guidance in their choices of courses, and questions concerning their competencies have been voiced in the Swedish media. All in all, the universities providing teacher education programs have been questioned in the evaluation study about the possibilities for future teachers both to get jobs, despite the expected shortage, and to be able to work as teachers, despite the fact that student control over courses and their future competencies are specified in the reform of 1999. The reports discussed on this article are directly related to the current debate in Sweden, since they concern aspects of the working conditions in an Information and Communication Technology (ICT)-supported distance-based teacher education program. Furthermore, this article has a more general value from an international perspective, since it concerns the possibility of providing for vocational education, such as teacher education, through a distance education mode and through programs that rest on assumptions of flexible delivery of learning experiences. Questions concerning flexibility and

choice as well as competencies and skills of prospective teachers become contextualised in the Swedish system, but at the same time they highlight problems at a more general level. However, in order to make the article understandable, it is necessary that a few words are said on the Swedish context and on the assumptions of the teacher education program in question.

Teacher Education in Sweden

Commissioned by the Swedish State, the Swedish Teacher Education Committee presented in 1999 a proposal in the Swedish Government Official Report Series for teacher education reform (SOU 1999, p.63). In the proposal the Swedish society is described as being in a phase of transition, in which unpredictable changes occur quickly. Teacher education must, as a consequence, also change. Teacher education today is supposed to give teacher trainees considerable freedom to create their own individual teacher education programs by choosing from the courses available. This choice seems to have two purposes. The Swedish State should both have control over certain common aims and objectives and at the same time allow each university enough freedom for local diversity and difference. Individual teacher trainees should have the possibility to develop unique competencies in order to be adequately prepared to meet a flexible and dynamic future society. This form of organisation also allows possibilities to adjust in a flexible way the program in accordance with local needs, something which applies to on-campus as well as off-campus education, in this case distance-based teacher education.

ICT-supported Distance-based Teacher Education

During recent years the rapid development and increased use of ICT has led to new options within all Swedish distance education programs. The flexibility in distance education has been presented as an example of how citizens can be given greater possibilities to educate themselves. The Swedish Distance Education Committee (SOU 1998, p.57; SOU 1998, p.83; SOU 1998, p.84) states that the aims of flexible learning include placing the control over where, when and how studies are to be conducted as far as possible with the students. This means that semesters, schedules and other forms of teacher control can be questioned and education can be undertaken and provided from different geographical locations. The interaction between the students and their teachers should be adapted to different conditions, to different places, times and to different media. Distance-based teacher education could then accommodate a number of different needs that individual teacher trainees might have. In the proposals it is emphasised that a significant degree of responsibility has to be placed on the individual. It is stated that no student can be forced to learn, and that it is the student who intends to study who should, in principle, decide by choosing between different alternatives. Consequently, it is up to the individual to take responsibility for making the right choices (SOU 1998, p.84).

Communication in distance-based teacher education has been increasingly conducted through the use of web-based learning environments. With ICT used both to conduct seminars and transmit lectures, an even higher degree of flexibility regarding geographical location and numbers of on-campus gatherings is assumed to be achieved. The increased use of computer-mediated communication (CMC) is, according to the rhetoric in the proposals (SOU 1998, p.83; SOU 1998, p.84), a contributing factor to this development. ICT is claimed to have the possibility to transform distance education that is based on the assumptions of flexible learning through individual studies at a distance and membership of a learning community. Thus the social aspects of education could influence the practice of distance-based education, which in relation to a vocational program such as teacher education would seem highly appropriate. A question arises, however, concerned with how the social aspect of education should be understood?

Emphasising a Social Aspect of Education – The Learning Community

Over the last decade educational theory has increasingly focused upon the social aspects of education. Learning is, accordingly, claimed to occur in collaboration (for example, Dillenbourg, Baker, Blaye and O'Malley, 1995) with others in some kind of social situation or context (for example, Lave, 1997), which education programs have to construct or reconstruct. Wenger (1998) has elaborated upon this in the concept of Community of Practice (CoP). According to Wenger (1998) a community could be defined as “a way of talking about the social configurations in which our enterprises are defined as worth pursuing and our participation is recognisable as competence” (p. 5). The learner must be given a possibility to participate in a community in order to create meaning and understanding through a process of negotiation that has inherent social, relational and temporal aspects. Practice is according to Wenger (1998) “a way of talking about the shared historical and social resources, frameworks, and perspectives that can sustain mutual engagement in action.” (p. 5). Participation in social practices shapes both the experiences of the human beings and the nature of the communities in which the experiences are drawn.

As a consequence of the development of the internet, distance education can today be understood as an interactive learning experience supported by the use of ICT (for example, Vrasidas and Glass, 2002). This shift implies that the concept of a learning environment can be considered to include a virtual dimension that laid down the ground for a web-based learning environment. This shift also marks the change from the use of ICT in educational settings as a tool for downloading educational material towards enhancing participation with others in a social context (for example, Bonk and Cunningham, 1998; Haythornwaite, 2002). Participation, both in a synchronic and asynchronous mode (for example, Kowch and Schwier, 1998), is enabled through web-based learning environments (for example, Bonk, 1998; Stephenson, 2001).

In discussions focused upon web-based learning environments, providing working conditions that enable participation in a social context is often emphasised, and it has also recently become a question in research on how such learning environments can foster the building and upholding of Online Learning Communities (OLC) (for example, Carlén and Jobring, in press; Haythornwaite, 2002; Lock, 2002; Seufert, Lechner, and Stanoevska, 2002). The community is here considered to be located on the internet and face-to-face meetings are reduced or even non-existence. One important feature in the OLC, concerns the creation of feelings of belonging to the community (cf. Palloff and Pratt, 2003) and how to distinguish membership among the members of the community (Haythornthwaite, Kazmer, Robins and Shoemaker, 2000). In the development of a joint group identity, characteristics are based on negotiated meanings integrated in to the practice of the community (for example, Schwier, Campbell and Kenny, 2004; Selznik, 1996). In other words, central to membership of an OLC is the sharing of history, ideas and values that create mutual engagements (cf. Wenger, 1998). These are all elements of importance when the concept of OLC is placed within the context of higher education and ‘recurrent education’ (to use a term proposed by Torsten Husén and advanced by Olof Palme) and particularly in relation to teacher education where the mutual engagement can be said to be directed towards the practice of teaching.

The educational framework considered in this article is an ICT-supported distance-based teacher education program in Sweden organised in terms of a flexible learning mode within the field of higher education and recurrent education. The aim, then, is to present a possible understanding of the working conditions, specifically understood as ‘working method’s, ‘examination forms’, and ‘students’ influence over their studies’, and to examine this understanding within the context of national steering documents governing the teacher education and theories emphasise the social aspects of education.

METHOD

In this study, data are analysed using a hermeneutical approach (for example, Gadamer, 1988, Ricoeur, 1995; Vattimo, 1997), with interpretations brought about by the use of a particular theoretical frame of reference. This approach is used, mainly based on beliefs that there are no objective facts to discover, instead research is based on interpretations that involve dimensions of uncertainty, which are hard to separate from each interpreter's own prejudices concerned with the issue under consideration (for example, Gadamer, 1988). The use of a theoretical frame of reference is inspired by Held (1996) who claims

For the framework we bring to the process of interpretation determines what we 'see', what we notice and register as important. Accordingly, particular interpretations cannot be regarded as *the* correct and final understanding of a phenomenon; the meaning of a phenomenon is always open to future interpretations from new perspectives. Interpretations are, therefore, always open to challenge. (p. 9)

In the following sections the data collection and procedures employed, participants and context, and the theoretical analytic tool are described.

Participants and Context

Participants were teacher trainees attending a Swedish teacher education course organised as a distance-based program in which ICT was used for administration, communication and support. The program under investigation was three and a half to four and a half years long, depending upon the teacher trainees' choice of examination. In the program, a web-based learning environment (WebCT) was used. The education provided was intended to be independent of space and time, and only to a minor degree dependant on physical meetings, called gatherings. The number of gatherings on campus differed from 1 to 4 per semester depending on semesters and courses. Teacher trainees worked in smaller groups, sometimes with a tutor from the university who, together with the teachers in each course, assisted the teacher trainees through their education. The number of teacher trainees enrolled at the time of data gathering was 77. The whole group was asked to participate and a total of 55 teacher trainees did so, thus approximately 71 per cent of this class of teacher trainees also worked together in smaller groups, with occasional help from tutors.

Data

Data were collected by the use of a questionnaire. The intentions in the construction of the questionnaire was to include areas or activities that could be understood as common and central features of the program under investigation, in order to supply data that provided an understanding of the working conditions involved in the program. The questionnaire contained different claims or questions to which answers were to be given in five categories. In the construction phase, the questionnaire was tested on other teacher trainees (n=23), and on university colleagues (n=5). After testing, the categories were adjusted and the questionnaire was altered not to include open questions. Depending upon the theme involved, the categories were 'never', 'seldom', 'a few times', 'often', and 'always', or, 'do not agree at all', 'to some extent do not agree', 'don't know', 'to some extent agree', and 'totally agree'.

Procedure

The questionnaire was published on the internet and all teacher trainees were approached through their web-based learning environment with a question about their participation. The questionnaire was reached through a hyper-link, protected by a password. Before the deadline for answering, a reminder was sent. After the deadline had expired, ordinary mails encouraging the teacher

trainees to answer within a week were sent to those who had not already filled in the questionnaire.

The Approach to Data Analysis

The procedure employed for the analysis of the data was built upon two dimensions with a bearing on education, conceptualised as a field of inquiry shown in Figure 1. The two dimensions of the field of inquiry involved two philosophical questions about humans, namely, a) the origin of human becoming, and b) the possibility of human agency. The questions were formulated in the following words.

1. Are humans individually or socially conceived?
2. To what extent do humans have possibilities to act and affect what he or she is becoming?

These two questions were captured in the field as the dimensions, individual – social and pre-determined – uncertain. The field enabled interpretations of how the working conditions in the program stressed these questions in different ways.

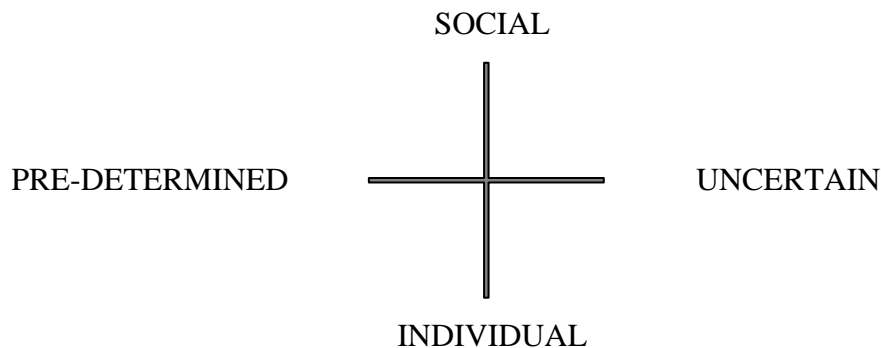


Figure 1. The two-dimensional field used in the analysis

ANALYSIS AND RESULTS

The three themes included here contained a total of 16 questions. The themes were: ‘Working methods’, ‘Examination forms’, and ‘Students’ influence over their studies’. Answers to the questions and the frequency of response categories for each question within the themes are presented in Tables 1, 2 and 3, where $N=77$ and $n=55$. The data used in the analysis is deliberately recorded in a form where neither the mean nor median value, nor any other statistical measure is used. The reasons for doing so are that the response patterns for each alternative are to be understood in relation to each other. Our focus is not, even if that were possible within other approaches, to conduct the analysis by using, for example, statistical significance.

Working Methods

The first theme is concerned with working methods, and it was examined using five different questions. The questions were intended to cover working methods believed to be common in the program investigated. Answers and frequencies are presented in Table 1.

In the answers regarding working methods, the two most common methods according to the teacher trainees are what can be called traditional forms of organising education. The most frequently used form, which 87 per cent of the teacher trainees describes as ‘often’ or ‘always occurring’, is the teacher talking and asking questions to which a single teacher trainee answer. This is followed by working in groups on assignments, which 79 per cent of the trainees describe as occurring ‘often’ or ‘always’. Slightly less common is that the teacher trainees’ works on their

own with the same assignments, since 69 per cent describe this as occurring ‘often’ or ‘always’. Teacher trainees also seem to work on their own conducting investigations, not as often as in the previous question, but ‘fairly often’ according to 56 per cent. Least common seems to be discussions between the teacher trainees and the teacher, ‘often’ or ‘always’ according to 47 per cent of the trainees. The teacher trainees seem to be left much on their own, with part of their work done in collaboration with other teacher trainees, and the teachers seem to rely on a high degree of teacher control when interacting with the teacher trainees.

Table 1. Frequency of ratings (%) in different categories on the questions in the theme ‘Working methods’

Questions	Never	Seldom	A few times	Often	Always
1. The teacher talks and asks questions, single students answer	0	2	11	69	18
2. The teacher and the students discuss together	0	13	40	45	2
3. Students work in groups on assignments	0	4	18	75	4
4. Students work on their own with the same assignments	2	7	22	67	2
5. Students conduct investigations of their own	0	7	36	56	0

When these results are analysed in relation to this field, the tendency by teachers to control the education is made explicit in that the most common way of working is when the teacher talks and asks questions to which a single teacher trainee answers. Consequently, the working methods still seem to be built upon a metaphor of transmission, from teacher educator to teacher trainee. Pre-determined working methods with a strong emphasis on the individual appear to be frequent, while those that imply social oriented aspects characterised by uncertainty seem to occur less often.

Examination Forms

The second theme concerned examination forms and contained seven questions. Each question gave an example of one examination form and when combined they were considered to be forms used in different degrees in the program investigated. Responses and their frequencies are presented in Table 2.

Table 2. Frequency of ratings (%) in different categories on the questions in the theme ‘Examination forms’

Questions	Never	Seldom	A few times	Often	Always
6. Oral examination	20	29	42	7	2
7. Written examination	0	7	33	49	11
8. Group examinations	0	9	45	44	2
9. Written home assignment	0	2	24	65	9
10. Accounts of individual work	0	22	35	36	7
11. Discussions in the study group	7	11	49	29	4
12. Drama, film or likewise	4	60	35	2	0

On the question of form of examination, the main form used seems to be relying on written text. Oral examinations, discussions and drama and other forms are reported as the least common. Written home assignment and written examination task, in terms of ‘often’ or ‘always’ occurring in the program is described with the percentage of 74 and 60 respectively. What is also apparent is the high degree of individual examination forms, group examinations, oral examinations and discussions in the study group only occur a few times according to almost half of the teacher trainees. Drama, film or likewise as well as oral examinations have 64 per cent and 49 per cent respectively occur ‘never’ or ‘seldom’.

When the results concerning examination forms are related to the field of inquiry, the same pattern evident for working methods is apparent. Often individual teacher trainees are examined

in ways in which the examination is based on written text. Tasks that are less conventional and uncertain forms of presentation, for example, drama or informal discussions together with teacher trainees on the same course seem to occur 'seldom'. This implies a strong emphasis on more traditional forms of examining individual teacher trainees by using relatively pre-determined methods of examination, that is, individual examination seem to occur most often.

Students' Influence over Their Studies

The third theme was concerned with teacher trainee influence and involved four questions. The questions concerned both influence over informal as well as formal working conditions. The responses and their frequencies are presented in Table 3.

Table 3. Frequency of ratings (%) in different categories on the questions in the theme 'Students influence over their studies'

Questions	Do not agree at all	To some extent do not agree	Don't know	To some extent agree	Totally agree
13. We students decide who to work with in group exercises	2	29	2	55	13
14. How we students should work with assignments is decided by us	4	35	2	56	4
15. We students decide the number of examinations	96	4	0	0	0
16. We students can decide what literature is used in the program	93	5	2	0	0

On the question of influence, it is apparent that the teacher trainees have most influence over the informal parts of their courses in teacher education, like regarding who they work with. As many as 68 per cent and 60 per cent respectively agree with this 'totally' or 'to some extent'. With respect to what content to work on and how the content is examined, the trainees there seem to have almost no influence at all. Thus 98 per cent of the teacher trainees responded that they 'totally' or 'to some extent agree' to have no influence over the literature used, and 100 per cent 'totally agree' or 'to some extent agree' that they have no influence over the number of examinations. There is at least more than half of the teacher trainees claiming to have some influence over the informal parts of their program of study, but most noticeable is the large number of teacher trainees who feel that only parts of these aspects are open to their influence. For instance almost 40 per cent feel that there is no room for influence on how they work on assignments.

When these results are related to the field of inquiry, both a similar pattern and a different pattern to the responses in the previous sections emerge. Similarities are the high teacher control and pre-determined content, with little or no possibility of influence given to the teacher trainees regarding literature and examinations. Differences seem to be that some degree of choice and possibility of influence is given on working methods and content in the group work. Socially oriented features are thereby given some degree of influence, being less controlled and more uncertain. These features are, on the other hand, given less attention by teacher education through the nature of assessment. The focus of the teacher education program is on individuals who manage pre-determined content.

DISCUSSION

In this article the aim is to present, from a learner-centered perspective, an understanding of the working conditions in a Swedish ICT-supported distance-based teacher education program, specifically with respect to 'working methods', 'examination forms', and 'students' influence over their studies'.

The analysis is directed toward an understanding in which individual responsibility and choice is promoted, while shared meaning combined with a collective and learner-centered responsibility is counteracted by a strong normative approach governing the organisation of the program. However, strong elements of individualism are present in the guiding rhetoric through SOU 1999, p.63. In the analysis it is possible to understand individualism as one aspect of how the working conditions, used in the program, foster the teacher trainees' understanding of learning and education. For example, the educational content seems to a considerable degree already to be pre-determined by the educational organisers, and the examination forms used tend to focus on the individual. The work done in groups is less frequently used for examinations but the teacher trainees have some influence over who to work with and how they can work. A paradox that implies both a freedom to choose and the creation of individual competence just as long as it is in line with the educational organisers' interpretations of adequate competence, that is defined by the Swedish State.

This line of reasoning is also plausible when looking at group work as a working method. One interpretation of why this working method is used could be that the teacher trainees are expected to lay the foundation for teacher-teams in their future workplace. The program lays the foundations for forming a future learning community of teachers. In SOU (1999, p.63) the idea was expressed as "the foundation for co-operation in the future profession is laid" (p. 126) and with respect to common values this was expressed as "that the teacher trainees can develop a common view regarding the learning and socialisation of children, youth and adults" (p. 127). A view more in line with thoughts of learning as a fundamentally social phenomena (for example, Dillenbourg et al, 1995) and the importance of that type of learning process includes participation in a social context, for example, a community (for example, Wenger, 1998). If so the sharing and negotiating of meaning and values became a focus. Group work is, though, a working method rarely used as a foundation for assessment and therefore the formation of a community within the program seems to be nothing else than wishful thinking.

The lack of social aspects in the program can be understood as if the distance-based teacher education, organised in a flexible learning mode, counteracts an integration of social aspects in terms of a common view of learning and socialisation. Social aspects, for example such as those expressed by Lave (1997) and Wenger (1998), where important activities for teacher trainees should be to have ongoing dialogues and discussions concerning both individual choices of the topics within the program and the educational content of these topics. In order to realise this in a program given at a distance, the educational organiser have to use ICT actively. Providing a web-based learning environment that fosters the building and the upholding of online learning communities where the teacher trainees meet virtually accommodates these intentions. The analysis implies, though, that the teacher trainees are not given many, or even any, opportunities to develop a joint group identity characterised by negotiated meanings integrated in the practice of a community (for example, Schwier, Campbell and Kenny, 2004; Selznick, 1996). The web-based learning used does not seem to fulfil its purposes, instead the limited number of gatherings on-campus each semester seems to be more or less the only time the teacher trainees can have an opportunity to develop a community characterised by shared history, enhancing of collaboration and meaning-making among the teacher trainees and thereby the creation of a growing mutual engagement for teacher education *per se* and their future practice as teachers (cf. Selznick, 1996; Wenger, 1998).

CONCLUDING REMARKS

In conclusion, it seems that the working conditions within the program reduce the possibilities to build and maintain online learning communities that can foster a shared history and experience among the teacher trainees. Whether this is due to the flexible learning mode and its lack of social aspects on learning and education, or due to the teacher educators understanding of how human

beings learns or due to national steering documents and the establishment of individualism is beyond the scope of this article. It seems, though, that participation in an online learning community has inherent within it, the possibilities to develop common value systems that apply both to more general questions, that involve an understanding of the world of human beings, as well as more particular questions that concerns the teacher trainees' future work in Swedish schools. To work as a teacher means to embrace certain values that operate as a kind of embodied guide for how the everyday work in schools both can and need to be conducted. These aspects do not seem to be present in the program investigated.

FUTURE RESEARCH

A most interesting research question to focus upon in a future study is how to include the ethical and morally oriented features of becoming a teacher in ICT-supported distance-based teacher education programs. This research is probably going to include, not only the questions addressed in this article, but also a more deepened scrutinising of the concept of an online learning community within teacher education.

REFERENCES

- Bonk, J.C. (ed) (1998). *Electronic Collaborators. Learner-Centered Technologies for Literacy, Apprenticeship, and Discourse*, Mahwah NJ, Lawrence Erlbaum.
- Bonk, J.C. and Cunningham, D.J. (1998). Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. In C.J. Bonk, (ed) *Electronic Collaborators. Learner-Centered Technologies for Literacy, Apprenticeship, and Discourse*. Mahwah, NJ: Lawrence Erlbaum.
- Carlén, U. and Jobring, O. (in press). The rationale of online learning communities. *International Journal of Web Based Communities*.
- Dillenbourg, P, Baker, M, Blaye, A. and O'Malley, C. (1995). The evolution on collaborative learning. In P. Reimann and H. Spada, (eds) *Learning in Humans and Machines. Towards an Interdisciplinary Learning Science*. London: Pergamon.
- Gadamer, H-G. (1988). *Truth and Method*. London: Sheed and Ward.
- Haythornwaite, C. (2002). Building social networks via computer networks – creating and sustaining distributed learning communities. In K.A. Renninger and W. Shumar, (eds) *Building Virtual Communities – Learning and Change in Cyberspace*. Cambridge: Cambridge University Press.
- Haythornwaite, C., Kazmer, M.M., Robins, J. and Shoemaker, S. (2000). Community development among distance learners: Temporal and technological dimensions. *Journal of Computer-Mediated Communication*, 6 (1), [Online] <http://www.ascusc.org/jcmc/vol6/issue1/haythornthwaite.html> [2005, January 25].
- Held, D. (1996). *Models of Democracy*. Cambridge: Polity Press.
- Lave, J. (1997). The culture of acquisition and the practice of understanding. In D. Kirshner and J.A. Whitson, (eds) *Situated Cognition: Social, Semiotic, and Psychological Perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Lock, J.V. (2002). Laying the groundwork for the development of learning communities with online courses. *Quarterly Review of Distance Education*, 3, 395-408.
- Kowch, E. and Schwier, R. A. (1998). Considerations in the construction of technology based virtual learning communities. *Canadian Journal of Educational Communication*, 26 (1), 1-12.

- Palloff, R. M. and Pratt, K. (2003). *The Virtual Student: A Profile and Guide to Working with Online Learners*. San Francisco, CA: Jossey-Bass.
- Ricoeur, P. (1995). *Hermeneutics and the Human Sciences* (J.B. Thompson, Trans.). Cambridge: Cambridge University Press.
- Schwier, R.A., Campbell, K. and Kenny, R. (2004). Instructional designers' observations about identity, communities of practice and change agency. *Australasian Journal of Educational Technology*, 20 (1), 69-100.
- Selznick, P. (1996). In search of community. In G. Angling, (ed) *Critical Issues in Instructional Technology*. New Haven: Yale University Press.
- Seufert, S., Lechner, U. and Stanoevska, K. (2002). A reference model for online learning communities. *International Journal on E-Learning*, 1, 43-54.
- SOU 1998:57 *Utvärdering av Distansutbildningsprojekt med IT-stöd* [Evaluation of the Distance Education Project Supported by IT], Stockholm: Utbildningsdepartementet.
- SOU 1998:83 *På distans – Utbildning, Undervisning och Lärande. Kostnadseffektiv Distansutbildning* [At a Distance – Education, Teaching and Learning. Cost-effective Distance Education; in Swedish]. Stockholm: Utbildningsdepartementet.
- SOU 1998:84 *Flexibel Utbildning på Distans* [Flexible Education at a Distance; in Swedish]. Stockholm: Utbildningsdepartementet.
- SOU 1999:63 *Att Lära och Leda. En Lärarutbildning för Samverkan och Utveckling* [To Learn and to Lead. A Teacher Education for Cooperation and Development; in Swedish]. Stockholm: Utbildningsdepartementet.
- Stephenson, J. (ed) (2001). *Teaching & Learning Online. Pedagogies for New Technologies*. London: Kogan Page.
- Vattimo, G. (1997). *Beyond Interpretation: the Meaning of Hermeneutics for Philosophy*. Stanford, CA: Stanford University.
- Vrasidas, C. and Glass, V.S (eds) (2002). *Current Perspectives on Applied Information Technologies: Distance Education and Distributed Learning*. Greenwich, CT: Information Age Publishing.
- Wenger, E. (1998). *Communities of Practice. Learning, Meaning and Identity*. New York: Cambridge University Press.

State educational policy and curriculum: The case of Palestinian Arabs in Israel

Ismael Abu-Saad

Ben-Gurion University of the Negev abusaad@bgu.ac.il

The state educational system in Israel reflects the declared character of the state as a 'Jewish state', and, consequently, the deep inter-group divisions in society, including a large Palestinian Arab minority. This study demonstrates how Israeli educational policy and curriculum are designed to support the Jewish nation-building project. As such, they silence the Palestinian Arab narrative while reshaping regional history for both Jewish and Arab students to fit the Zionist narrative. Furthermore, Israeli educational policy has played an essential role in consigning Palestinian Arabs to the social, economic and political margins of Israeli society.

Educational policy, curriculum, Palestinian Arabs, Israel, minorities

Israeli society is very heterogeneous due to the existence of a wide variety of immigrant Jewish populations and an indigenous Palestinian Arab population; yet, at the same time, is officially defined as 'the State of the Jewish people.' The centrality of the notion of 'Jewishness' to Israel's national identity has been translated, in practical terms, into the subordination of the indigenous Palestinian Arab minority, which comprises 19 per cent of the total population, to the Jewish majority (Abu-Saad, 2004a, Fares 2004; Lewin-Epstein and Semyonov, 1993).

This article examines the role of Israeli state educational policy and curriculum in this process. While the state educational system is administratively subdivided into a Jewish system (which is further divided into a number of subsystems, for example secular schools, religious schools), and an Arab system, this reflects not only the cultural and language diversity, but also the deep inter-group divisions in society. Numerous studies have documented the differential allocation of resources to these systems, and the poorer educational outcomes in the Arab system, as measured by drop out and matriculation rates, and proportion of age cohorts accessing higher educational opportunities (Abu-Saad, 2004a; Golan-Agnon, 2005; Human Rights Watch, 2001). Furthermore, the article focuses upon the question of how Israeli educational policy positions Palestinian Arab citizens of Israel within the national framework, and what implications these educational policy choices have for the broader society.

THE CONTESTED HISTORY OF 'BELONGING'

In the late 1800s, the Zionist nationalist movement was developed by a group of Jewish intelligentsia in Europe, the goal of which was to establish a Jewish state in Palestine. Zionism was based on the premise that Palestine was a territory that belonged exclusively to the Jewish people due to their presence on the land during biblical times. The Zionist movement portrayed Palestine as a 'land without a people, for a people without a land' (Masalha, 1997), and the Zionist immigrants to Palestine as pioneers coming to conquer an inhospitable environment, and make the barren desert bloom. The notion of an empty territory was used to justify Zionist colonisation, with its dehumanising orientation toward the native population, leading to their delegitimation as a people, belonging to that particular place (Masalha, 1997). From its inception, the Zionist movement sought out and gained the support of the era's major European

colonial powers, and most notably Great Britain. In 1947, the United Nations Organisation (UN) voted to partition Palestine into two states, one Jewish over 56 per cent of the territory (in which Jews comprised 51 per cent of the population and owned 10 per cent of the land), and one Arab, over 42 per cent of the territory; despite the fact that Palestinian Arabs represented over 67 per cent of the total population of Palestine at that time (Hadawi, 1991; Lustick, 1980). The indigenous Palestinian Arabs rejected the partition plan, and as soon as the British withdrew, turning the unresolved conflict over to the UN in 1948, the Zionist leadership declared Israel's independence as a 'Jewish' and 'democratic' state. The Arab states declared war on the Jewish state, but the Israeli forces defeated the Arabs, and by the time of the ceasefire, had taken control of 77 per cent of the land of Palestine.

AIMS AND GOALS OF PUBLIC EDUCATION IN ISRAEL

Many nations that consider themselves liberal democracies, and have indigenous and minority populations under their jurisdictions, are advertently or inadvertently monocultural in the formulation and provision of educational services, and an analysis of Israeli educational policy indicates that this has generally been the approach in Israeli public education. Private Zionist/Jewish schooling was well-developed in Palestine prior to the establishment of the Israeli state in 1948, and after its establishment, it seemingly continued to envision and to develop its educational system as if it were still only educating Jews. Israel's *1953 Law of State Education* specified the following aims for the education system:

to base education on the values of Jewish culture and the achievements of science, on love of the homeland and loyalty to the state and the Jewish people, on practice in agricultural work and handicraft, on pioneer training and on striving for a society built on freedom, equality, tolerance, mutual assistance, and love of mankind. (Mar'i, 1978, p. 50)

Over 50 years have passed since the enactment of this law, but the aims it specified remain central to current Israeli public educational policy. Though the law was amended in 2000, it maintains educational objectives for public schools that emphasise Jewish values, history and culture, while ignoring Palestinian values, history and culture (Adalah, 2003). These narrowly-defined educational aims that speak to the identity of three-fourths of the state's students while overlooking the other fourth, have continually been reaffirmed in the official discourse about education in Israel. In June 2001, Minister of Education, Limor Livnat, stated that she would like to see that "there is not a single child in Israel who doesn't learn the basics of Jewish and Zionist knowledge and values" (Fisher-Ilan, 2001, p. 4B). The Ministry of Education operationalised these goals through programs such as the "100 Basic Concepts" curriculum unit that was introduced to the middle schools in the 2004/05 school year (Ministry of Education and Culture, 2004). While separate lists of the 100 key concepts were developed for the Jewish and Arab education systems, they largely reaffirmed the subordinate and superfluous status of Palestinian Arabs in Israel. One third of the concepts were devoted to heritage, and the list for the Jewish school system was entitled, "Concepts in Jewish Heritage", while the list for the Arab school system was entitled, "Concepts in Arab Heritage for the Arabic Sector", a qualification suggesting that they were of no importance or relevance for any other sector of Israeli society.

The 34-item Jewish list was comprised of broad concepts about ancient Jewish history and religion, and national holidays (for example, Purim, Independence Day, Hanukah, Jerusalem Day, which despite being called 'national' are not holidays for all citizens of the country). It also included broader social concepts, such as respect for parents and teachers, as a part of Jewish heritage. The 34-item Arab list contained concepts from both the Muslim and Christian religions, thus providing a more superficial treatment of each; and other general concepts chosen as characterising the Arabic culture from a perspective of romanticising the orient (for example,

Arab markets, hospitality, generosity, the tent). At the same time, it excluded the broader social concepts included in the Jewish heritage list (for example, respect for parents and teachers), as though such values were unique to Jewish culture and not present in Arab culture.

The second list in the 100 Basic Concepts program for the Jewish schools was entitled, "Zionist Concepts." It included 33 items dealing with the Zionist movement, 15 prominent, modern Zionist or Israeli leaders (including 3 women), the "wars of Israel", pre-and post-state waves of immigration, and institutions that have become inseparable and indistinguishable from Zionism, such as the Holocaust Museum and the Israeli military. The parallel 33-item list for the Arab schools was entitled, "Zionist Concepts for the Arab Sector" and included the same concepts as the Jewish list, with a few exceptions. The Arab list included the names of three Arab citizens of Israel (one political figure from the mainstream Zionist Labor Party, one novelist and the only Arab ever to receive the Israel Prize, and one Christian religious figure, all of whom were men). It also included a center for Arab-Jewish Culture, Society, Youth and Sports (House of the Vine) that organises non-political social meetings between Arabs and Jews aimed at creating recognition and understanding, and educating for co-existence, good neighbouring and tolerance through cultural, artistic and community activities. Neither the three Arab names, nor the House of the Vine Center nurturing Arab-Jewish coexistence, appeared on the Jewish list of Zionist concepts, which instead included additional items about the pre-state Zionist settlers and their victories over the indigenous population, the unauthorised pre-state immigration of Jews to Palestine, and the absorption of the first massive post-state wave of Jewish immigration. Not surprisingly, there was not a single mention for either Jewish or Arab students of the history of the Palestinian people, the consequences they suffered (dispersion and dispossession) as a result of the fulfilment of Zionist aspirations through the establishment of Israel or the Palestinian national movement. In stark contrast, Palestinian Arab students were required, along with Jewish students, to **memorise** a substantial list of Zionist historical facts and figures.

The final section of the 100 Basic Concepts for both school systems was entitled, "Concepts in Israeli Democracy". It contained the same broad humanitarian items (for example, human rights, the Geneva Convention, Rights of the Child, pluralism, humanism) for both groups, in addition to less inclusive items such as that defining Israel as the "Jewish and Democratic State," the *Law of Return* (which applies to Jewish immigration and return rights only), the flag and the national anthem (which are both symbols of Jewish religious origin).

Thus, while the *1953 Law of State Education* might seem to represent an out-dated policy approach that would have changed over time to recognise and accommodate the presence of a large Palestinian Arab minority in the student body; in fact, a program such as the "100 Basic Concepts" demonstrates how the educational aims and goals that were established in the *1953 Law of State Education* have continually been renewed and reaffirmed, keeping Palestinian Arabs from being fully present in their own education, and basically absent from Jewish education.

The Arab educational system in Israel has been, and continues to be, governed by a set of political criteria which Palestinian Arabs have no say in formulating (Al-Haj, 1995; Mar'I, 1978; Swirski, 1999). While the *1953 Law of State Education* strongly emphasised the development of Jewish identity and values, no parallel aims were ever set forth for the education of Arabs in Israel, though in the 1970s and 1980s some attempts were made by committees directed by Jewish educators (Al-Haj, 1995). Nor was the Palestinian minority ever given autonomous control over their education system or allowed to determine its aims, goals and curricula. Though the Arab school system has a separate curriculum, it is designed and supervised by the Ministry of Education, where virtually no Arab educators or administrators have decision-making powers. Despite the fact that Arabic is the medium of instruction in the schools, the Arab school system does not represent multicultural recognition and accommodation, nor is it, in the words of Freeland (1996, p. 182), an example of "indigenous control over education and true

interculturality." A high-level Jewish administrator in the education system described the reality as follows: "...the Arab head of the Arab education system has no authority or budget, he never even says anything at the meetings. Between us, we call him 'the plant.' His deputy, a Jewish man appointed by the security service, actually runs the department" (Golan-Agnon 2005, p. 207).

This contrasts sharply with the state's Jewish religious school system, which is physically and administratively separate from the state's secular Jewish school system, and maintains completely autonomous control over its educational policy, aims and goals (Adalah, 2003; Mar'i, 1978; Swirski, 1999).

CURRICULUM IN THE JEWISH SCHOOL SYSTEM

Consistent with national educational goals, the curriculum in Jewish schools places a strong emphasis on the development of national identity, active belonging to the Jewish people, and furthering of Zionist aspirations. At the same time, there is little or no recognition of Arab history and Palestinian Arab history in particular. Where the curriculum includes reference to Arabs, it generally tends to take an Orientalist approach, portraying them and their culture in a negative light. Said analysed the way in which Eastern cultures are viewed, described and represented by Western academic scholarship, politics, and literature in his book *Orientalism* (1978) and numerous subsequent works. Said's main conclusions were aimed at how the Western (specifically the British, French, and more recently American and Israeli) economic, political and academic powers have developed a dichotomised discourse in which an inherently superior West was juxtaposed with an Eastern 'Other' according to terms and definitions determined by the West itself. Orientalism has created an image of the Orient as separate, backward, silently different, irrational and passive. It is characterised by despotism and resistance to progress; and since the Orient's value is judged in terms of, and in comparison to the West, it is always the 'Other', the conquerable and the inferior.

School textbooks are widely recognised as important agents of socialisation that transmit and disseminate societal knowledge, including representations of one's own and other groups (Bar-Tal and Teichman, 2005). According to Luke (1988), school textbooks "act as the interface between the officially state-adopted and sanctioned knowledge of the culture, and the learner. Like all texts, school textbooks remain potentially agents of mass enlightenment and/or social control" (p.69). As Down (1988) stated, textbooks dominate what students learn at school, and determine the curriculum, as well as the facts learned, in most subjects. In addition, the public tends to regard textbooks as essential, authoritative, and accurate knowledge, while in most school systems, teachers rely on them to organise lessons and structure subject matters. This is particularly true in Israel, since teachers are obliged to base their instruction upon Ministry of Education-approved textbooks. According to Bar-Tal and Teichman (2005):

Due to the centralised structure of the educational system in Israel, the Ministry of Education sets the guidelines for curricula development and has the authority to approve the school textbooks. Thus, the ministry outlines the didactic, scholastic and social objectives to be achieved (Eden, 1971), and the textbooks' contents reflect the knowledge that the dominant group of society is trying to impart to its members. (p. 159)

Bar-Tal and Teichman (2005) summarised the major studies done on textbooks in Jewish schools and the history of their reforms. Firer's (1985) study examined history textbooks between 1900 and 1984, and their role in promoting Zionist socialisation. Firer found that all of the history books in the pre-state period (1900-1948) stressed the exclusive rights of the Jewish people to ownership of Palestine. Arabs, in turn, were portrayed as a backward, primitive people with no similar ownership rights in the 'neglected' land that was awaiting 'Jewish redemption'. As violent conflict began to erupt due to the opposing nationalisms of the indigenous Palestinian Arabs and

the Zionist settlers, Jewish history textbooks also began to refer to Arabs undifferentiatedly as easily agitated robbers and vandals. Bar-Gal's (1993) study of geography textbooks in the same period produced similar findings. The earliest geography textbooks, produced by Zionist authors who lived in Europe and endorsed the view of 'a land without a people for a people without a land', tended to completely ignore the presence of the indigenous Arab population in Palestine. Later, the textbooks by authors living in Palestine were characterised by attitudes of ethnocentrism and superiority toward Arab society, similar to the common European attitudes toward the Orient at that time. As violent conflict with Palestinian Arabs erupted, they began to be represented as 'the enemy', and according to Bar-Gal (1993), were described as a:

...negative homogeneous mob that threatens, assaults, destroys, eradicates, burns and shoots, being agitated by haters of Israel, who strive to annihilate the most precious symbols of Zionism: vineyards, orange groves, orchards and forests. Again, the Arabs were viewed as ungrateful. According this view Zionism brought progress to the area and helped to overcome the desolation, and thus helped to advance also the Arabs. But instead of thanking the Jews for building the country for the benefit of all its citizens, they respond with destruction and ruin (p.181).

From the establishment of the state of Israel through the early 1970s, school textbooks continued to present Arabs negatively, according to the same ideological-educational perspective adopted during the pre-state period (Bar-Tal and Teichman 2005; Firer, 1985; Podeh, 2002). According to Firer (1985), the first textbooks published by the newly-founded state were influenced by the trauma of the Holocaust in Europe, and used the same emotive concepts from that experience to describe the Jewish-Arab conflict. As Bar-Tal and Teichman (2005) described, these textbooks completely removed the Jewish-Arab conflict from its actual context:

Most of these books did not even mention the existence of a Palestinian nation, never mind its aspirations or the driving forces behind Palestinian nationalism. Thus, the Arabs' violence and resistance to Zionism, presented without explanation, looked absolutely arbitrary and malicious. It interfered with the noble and peaceful attempts of the Jews (described as victims) to return to their homeland. (p. 162)

The critical omission of Palestinian Arabs' history, pre-1948 life in Palestine, national aspirations, and their consequent dispossession, was in actual fact the ultimate delegitimisation of their identity and struggle. It was central to the Zionist narrative of the history of "*Eretz Israel*" (Land of Israel), "as a land without a people for a people without land" and the ancient and external homeland of the Jews, that was disseminated through the Ministry of Education textbooks. The curriculum up until the late 1960s was concerned primarily with the needs of nation-building and the construction of a homogeneous national identity, and to this end, it used mechanisms of denial, omission and exclusion toward Arabs (Bar-Tal and Teichman 2005).

From the late 1970s to the early 1990s, a few textbooks began to include references to and descriptions of Palestinian nationalism, or positive descriptions of the Arab citizens of Israel (though still disregarding their Palestinian identity) written by Jewish Zionist writers and from a Zionist perspective, however, the use of these textbooks was optional and they were eventually dropped from the curriculum (Bar-Tal and Teichman 2005). The overwhelming trend in portraying Arabs during this period in history, geography, civics studies and Hebrew (readers) remained negative (Bar-Tal and Tichman 2005).

The Zionist historical narrative perpetuated the image of the Arab, and the Palestinian Arab in particular, as an ahistorical, irrational enemy. A 17-year-old Jewish high school student described the contents of the schoolbooks in Jewish schools and viewpoints expressed by some Jewish teachers as follows:

Our books basically tell us that everything the Jews do is fine and legitimate and Arabs are wrong and violent and are trying to exterminate us... We are accustomed to hearing the same thing, only one side of the story. They teach us that Israel became a state in 1948 and that the Arabs started a war. They don't mention what happened to the Arabs-they never mention anything about refugees or Arabs having to leave their towns and homes... Instead of tolerance and reconciliation, the books and some teachers' attitudes are increasing hatred for Arabs (Meehan 1999, p.20)

Furthermore, some of the geography textbooks for high school published in the early 1990s portrayed Arabs in terms of a 'demographic problem' – a new type of threat to the Zionist vision of a 'Jewish *and* democratic' state (Bar-Tal and Teichman 2005; Orni and Efrat, 1992; Sorkis, Raf and Sharar, 1991).

The curriculum in Jewish Israeli schools has been instrumental in explicitly and implicitly constructing racist and threatening stereotypes and a one-sided historical narrative that, through the educational system, is internalised in the Jewish Israeli psyche; and that has, in turn, provided the basis for maintaining a deeply divided society and its many discriminatory practices. As a former-Israeli academic, Oren Ben-Dor, stated about his educational experience in the Jewish school system:

All my education in Israel was one sided, treating the other [for example, the Arabs] as the enemy, the murderers, the rioters, the terrorists ... without alluding, in any way, to their pains and longings. For my teachers and, as a result, for me also, for many years, Zionism was beyond reproach; it was a return to the promised land as a result of persecution, it was draining the swamps, it was building a state based on Jewish genius. (Ben-Dor, 2005)

According to Podeh (2002), however, analysis of history textbooks for the higher grades published toward the end of the 1990s indicated a major and significant change in the depiction of Palestinians, Palestinian nationalism, Arabs, and the Israeli-Arab conflict. Some of these textbooks included recently declassified Israeli governmental archival materials and were based on critical historical research that shed a more balanced light on the conflict and for the first time portrayed Palestinian Arabs not only as spectators or aggressors but also as victims of the conflict. However, even with these much celebrated revisions in textbooks, Raz-Krakotzkin noted that:

...in all the textbooks there is not one single geographical map which shows the [pre-1948 Palestinian] Arab settlements – only the Jewish settlements are shown. Generally speaking, the land itself has no history of its own, and the history of the land is presented as the history of the Jewish myth about it. The whole period, between the second temple and the Zionist settlement is not taught at all. But more precisely, the Israeli student has no idea whatsoever about the settlement of the country before '48, that is to say, has no idea about the history of the expelled themselves and of their life before the expulsion. And so the mythical image of the country was created as 'the Promised Land of the Jews' and not as a cultural-geographical entity in which the [Jewish] colonisation took place. (1999, p. 5)

Even with the deficiencies Raz-Krakotzkin noted, though, the publication of the new history textbooks led to heated debates in Israeli society, and in November 2000, the parliamentary Education Committee decided to delay the use of one of these textbooks. Bar-Tal and Teichman (2005) characterised this as "an act that shows that part of the society and its representatives have difficulty in accepting changes in school textbooks that question the Zionist narrative" (pp. 72-3). They further suggested that this decision may have been due to a counter-trend in Israeli society brought about by the "outbreak of violence" with the Al-Aqsa Intifada in September 2000. They

concluded that since the parents and grandparents of the present generation were consistently presented with negative images of Arabs in school textbooks, it would take many years, indeed another several generations, to rewrite and introduce a balanced presentation of Arabs into the school textbooks, without negative stereotypes and delegitimising labels. However, according to their own argument, maintaining this same approach in the textbooks can only perpetuate the conflict as additional generations of Israeli Jews are educated in a manner that continues to deny the history, humanity, legitimate grievances, and aspirations of Palestinian Arabs. Paradoxically, it is precisely a more balanced picture of Palestinian history and aspirations that would enable the Israeli Jewish population to understand the roots of, and perhaps even begin to break out of the cycle of violence.

The Jewish school system further contributes to the marginalisation of the Palestinian minority by giving Jewish students little, if any, exposure to the Arabic language or culture, directly, rather than filtered through Zionist lenses. Despite the fact that Arabic is one of the two official languages in Israel, the study of Arabic is not required in Jewish schools as a matriculation subject (for example, a requirement for obtaining a high school diploma). Less than 4 per cent of Jewish high school students voluntarily study Arabic as one of their matriculation subjects (Lev-Ari, 2003). According to the Education Ministry Director General, Ronit Tirosh, Jewish students feel antagonistic toward the Arabic language. Tirosh stated that:

[Arabic] is a language that is identified with a population that makes your life difficult and endangers your security. Even so, students understand that knowing Arabic helps them to view life in Israel through the eyes of the Arabs...We thought about making Arabic compulsory for matriculation, but concluded that if less than 10 per cent of students learn it voluntarily, it would be impossible to force it on the rest. (Lev-Ari, 2003)

CURRICULUM IN THE ARAB SCHOOL SYSTEM

In sharp contrast to the promotion of a Jewish and Zionist identity in the curricular goals and materials in the Jewish schools, the curricular goals that the central Ministry of Education developed for Arab education tend to blur rather than enhance the formation of an Arab identity. Palestinian identity in particular is treated as something at best irrelevant and at worst, antithetical, to the overriding goals and aims of the Zionist educational project. Thus, consistent with the "100 Concepts" program, the aims of the Arab educational system, as well as specific curricular goals, require students to learn about Jewish values and culture, while receiving superficial exposure to carefully screened and censored Arabic values and culture, and the results of this can be seen clearly in the government-controlled curriculum for elementary and secondary schools (Abu-Saad, 2004b; Al-Haj, 1995; Mar'i 1978, 1985; Peres, Ehrlich and Yuval-Davis, 1970). Arab students are required to spend many class hours in the study of Jewish culture and history and the Hebrew language (and in total, more than they spend on Arabic literature and history). Thus, they are required to develop identification with Jewish values and further Zionist aspirations at the expense of the development of their own national awareness and sense of belonging to their own people. The Arab national identity is much less emphasised, and the Palestinian identity goes completely unrecognised (Al-Haj, 1995; Ma'ri, 1978; 1985). Furthermore, the basic goal of Jewish studies in Arab education is not the development of cultural competence as a bridge to Jewish Israeli society but is rather to make Arabs understand and sympathise with Jewish and Zionist causes and blur their own national identity in Israel (Al-Haj, 1995; Mar'i 1978, 1985; Swirski, 1999).

In the 1970s, a group of Jewish Israeli researchers, Peres, Ehrlich and Yuval-Davis, addressed the same issues. They criticised the curriculum imposed upon Arab schools by the Ministry of Education for attempting to instil patriotic sentiments in Arab students through the study of

Jewish history, and pointed out the absurdity of the orientalist expectation that the “Arab pupil ... serve the state not because the latter is important to *him* and fulfills *his* needs, but because it is important to the Jewish people” (Peres, Ehrlich, and Yuval-Davis, 1970, p. 151).

Nevertheless, the Arab educational system has maintained its emphasis on the Zionist national project that has dispossessed and continues to marginalised the Palestinian people; while at the same time suppressing the students’ knowledge of and identification with the Palestinian and broader Arab peoples/nations. Specifically, though Palestinian Arab students are required to read the literature and poetry of the Zionist movement, celebrating the establishment of the Jewish state in Palestine, their curriculum does not include the Palestinian Arab literary classics studied throughout the Arab world (Adalah, 2003). As a Palestinian Arab student stated:

Everything we study is about the Jews. Everything is Jewish culture. We study Bialik and Rachel [Jewish nationalist poets]. Why do I have to study them? Why don’t they teach me Mahmud Darwish [Palestinian nationalist poet]? Why don’t they teach me Nizar Qabbani [Arab nationalist poet]? Why don’t they teach me Edward Said? Why don’t they teach me about Arab philosophers and Palestinian poets? I know that my Arabic language is not very strong, because I know if I don’t speak fluent Hebrew I can’t function in this country.... I know that the Arabic language in Palestine is endangered. Schools, not individually, but the educational system as a whole has a very negative impact on our identity. The whole world now recognises the existence of Palestine and that there is something called the Palestinian people. So why are they still teaching me about Bialik and Rachel? What is the problem in teaching us Palestinian history? The problem is that they are afraid. They don’t want us, Palestinian Arabs, to develop an awareness of our national identity. (quoted in Makkawi, 2002, p.50)

This suppression of Arab identity, culture, and political concerns has incessantly been maintained in the curriculum for Arab schools. Consistent with the orientalist approach of imposing the so-called ‘superior’ Western (and in this case, Zionist Israeli) perspective, interpretations and priorities upon the ‘inferior Other,’ the curriculum for the Arab education system is designed to further and implement the aims of Zionism among the state’s Palestinian Arab students through the displacement of their history and nationality with the Zionist narrative.

In 1978, the late Arab educator and researcher, Sami Mar’i described the explicitly monocultural status of Palestinian Arab education within the Israeli public school system in the following terms which, unfortunately, still provide an accurate description over 25 years later:

Arab education is a victim of Israeli pluralism not only in that it is directed and managed by the majority, but it is also a tool by which the whole minority is manipulated....[It] is not only an example of the Israeli pluralism by which Arabs are denied power, it is also a means through which the lack of power can be maintained and perpetuated. Arab citizens are marginal, if not outsiders.... The Arab Education Department is directed by members of the Jewish majority, and curricula are decided upon by the authorities with little, if any, participation of Arabs. Arab participation does not exceed writing or translating books and materials according to carefully specified guidelines, nor does it extend beyond implementing the majority’s policies. (Mar’i, 1978, p. 180)

Reform efforts have repeatedly failed to bring about change, since none of the recommendations of the many committees appointed by the government to study or improve the Arab education system have ever had any binding power (Abu-Saad, 2001; Al-Haj, 1995). As such, Palestinian Arab students continue to be subjected to a curricular and educational program designed to

address the needs and meet the concerns of the ruling majority, and ensure the marginalisation and subordination of the minority.

SOCIO-ECONOMIC IMPACT OF ISRAELI STATE EDUCATIONAL POLICY

Current Israeli public statistical data reveals that the Palestinian Arab population has higher levels of unemployment (14 per cent versus the national average of 9 per cent), lower average income (3,992 NIS versus the national average of 6,314 NIS), and almost twice the rate of children living in poverty (58 per cent) as in Israeli society as a whole (31 per cent) (Fares, 2004). Since these national averages include the Arab sector, the above figures de-emphasise the extremity of the gap between the Israel's Palestinian Arab and Jewish citizens.

The government produces an official report grouping local authorities in Israel into socio-economic clusters. In its most recent report, the 82 Palestinian Arab Local Authorities in Israel were primarily located at the bottom of the scale. They represented an overwhelming majority in the lowest clusters, making up 80 per cent of the first (lowest), 93 per cent of the second, and 87 per cent of the third cluster, and had no representation at all in the highest four clusters (Fares, 2004).

Another key indicator of the marginalisation and subordination of the Palestinian Arab community is the glaring lack of parity in the hiring and promotion of Arabs within the nation's civil service. This is a concern for Palestinian Arab citizens because civil service jobs are a direct and indirect gateway to professional advancement in various realms of the public sector. In 2003, of total of 55,409 civil service employees nationally, only 2,798 (5 per cent) were Palestinian Arabs (Fares, 2004).

The marginalisation of Palestinian Arabs in Israel through political, economic and social discrimination, and their subordination to the Jewish majority in almost every aspect of stratification, including income, occupational distribution, employment participation, land ownership and community-level infrastructure and development has been extensively documented (Fares, 2004; Haider, 2005; Lewin-Epstein and Semyonov, 1993). Its implications for the Palestinian Arab citizens of Israel were articulated by Zeidani (2005), who stated:

The state of Israel discriminates against me negatively, it deprives and neglects me, consigns me to the economic, social and political margins.... The [Palestinian] Arabs in Israel are a more or less homogeneous ethnic/national group, a group that differs in a clear and distinct way from the Jewish majority. There are differences between us, barriers and boundaries of all types: a different language, a different religion, a different culture, and a different, if not rival, historical narrative.... We have a separate education system that is controlled by a handful of Jews, who decide its content and goals, and fill the system's various positions. In other words: we are different and separate, and because we are different and separate, we do not enjoy equal status, and therefore, we are pushed to the margins. This is to say, that this marginality is a consequence of coercion – and not a result of free choice or free will. The [Palestinian] Arabs in Israel are...struggling to escape from the coerced marginality; they want to be full and active participants in determining their future and their fate. (pp. 89, 91-92)

CONCLUSIONS

This study demonstrates how Israeli educational policy and curriculum are designed to serve the Zionist national project. As such, they perpetuate racist and hostile images of Arabs to Jewish students, and silence the Palestinian Arab narrative while reshaping regional history for both Jewish and Arab students to fit the Zionist narrative. While the sense of Palestinian Arab

belonging to the Zionist national project (for example, building the Jewish state) can only be partial and incomplete, if it exists at all, the development of identification with the Palestinian people and Arab peoples more broadly is suppressed. The study of extensive required curricular materials is used to make the Palestinian Arab student understand the history and empathise with the suffering of the Jewish people. Thus, the policy and content of the state-controlled education system for Palestinian Arabs aim to re-educate the students to accept the loss of their history and identity. And it prepares them, ideologically and practically, to accept the superior status of the Jewish people, and the subordination of their needs and identity to the needs of the national Zionist project.

The consequences of the aggressively monocultural educational approach the Israeli state has adopted, not only for the Palestinian minority, but also the Jewish majority, are grave. The stereotypical, negative and ahistorical picture of Palestinian Arabs fostered in the Jewish school system promotes racism and cultural imperialism rather than openness to cultural pluralism and democracy. No serious effort is made to provide Jewish students with the tools they need for cultural competence in Palestinian and Arab societies linguistically, culturally, or historically. Instead, their school system encourages them to maintain a sense of distance from and superiority over the Palestinian Arabs who are citizens of Israel. It also serves to cripple any efforts to resolve the conflict over land, nationality and the basic rights of Palestinian Arabs (whether those holding Israeli citizenship, living in the Occupied Palestinian Territories, or living as refugees) since they are portrayed as a non-people, without a history.

The state's educational policy and curriculum for the Arab school system on one level appears to provide Palestinian Arab students with the tools they need for attaining cultural competence in Jewish Israeli society through extensive historical, linguistic, and socio-cultural studies of the Jewish people and Hebrew language and literature. Ironically, however, the system's concurrent efforts to re-educate the Palestinian students to forget their own history and identity, coupled with its discriminatory practices, seem to end up reinforcing the students' Palestinian identity and sense of the conflict with the Jewish majority. Taylor (1994) commented that "...misrecognition shows not just a lack of due respect. It can inflict a grievous wound, saddling its victims with a crippling self-hatred. Due recognition is not just a courtesy we owe people. It is a vital human need" (p.26). This study suggests that misrecognition can also create resistance among minority group members to learning about the majority's culture, while majority educational policy is at the same time ignoring, actively suppressing or trivialising the study of the minority's culture. Though the state-sponsored curriculum in Arab schools is aimed at suppressing and erasing the Palestinians' collective memory, it seems instead to succeed at providing them such an alienating educational experience that it fosters greater bitterness, enmity and separation between Israel's Jewish and Palestinian Arab communities.

One can only question whether this situation of discriminatory and antagonistic separation is, indeed, in the long-term interests of the State, whose ideology and mythology notwithstanding, is in fact a multi-ethnic state, with an indigenous minority that makes up nearly one fifth of the population. For the present, the situation seems to be satisfactory to the Jewish majority, and the public education system will continue to aid in perpetuating it, with considerable impact. However, as the sense of bitterness and alienation grows within the Palestinian Arab population, so does the threat of political and civil instability.

REFERENCES

- Abu-Saad, I. (2004a). Separate and Unequal: The Consequences of Racism and Discrimination against Palestinian Arabs in the Educational System in Israel, *Social Identities*, 10 (2), 101-127.

- Abu-Saad, I. (2004a). The Role of Education in Minority Identity Formation in an Ethnic State. *Arab Studies Journal*, Vol.XI (1) Vol. XII (2), p.57-75.
- Abu-Saad, I. (2001). Education as a tool for control vs. development among indigenous peoples: The Case of Bedouin Arabs in Israel. *Hagar: International Social Science Review*, 2 (2), 241-259.
- Adalah Report (2003). *Education Rights – Palestinian Citizens of Israel*. Shafa'amr: Adalah the Legal Center for Arab Minority Rights in Israel.
- Al-Haj, M. (1995). *Education, Empowerment and Control: The Case of the Arabs in Israel*. State University of New York: Albany, NY.
- Bar-Gal, Y. (1993). *Moledet and Geography in one Hundred Years of Zionist Education*. Tel Aviv: Am Oved. (in Hebrew).
- Bar-Tal, D. and Teichman, Y. (2005). *Stereotypes and Prejudice in Conflict: Representations of Arabs in Israeli Jewish Society*. Cambridge: Cambridge University Press.
- Ben-Dor, O. (2005). Academic freedom in Israel is central to resolving the conflict, *Counterpunch*, May 21-22. [Online] www.counterpunch.or/bendor05212005.html [18 March 2006].
- Down, A. (1988). Preface. In H. Tyson-Bernstein, *A Conspiracy of Good intentions: America's Textbook Fiasco*. Washington, DC: Council for Basic Education.
- Eden, S.1(971). *On the new Curricula*. Jerusalem: Maalot. (in Hebrew).
- Fares, A. (2004). *The State Budget and the Arab Citizens in Israel*, Haifa: Mossawa Center. [Online] <http://www.mossawacenter.org/he/reports/2004/12/041224.html> [18 March 2006].
- Firer, R. (1985). *The Agents of Zionist Education*. Tel Aviv: Hakibutz Hameuhad. (in Hebrew).
- Fisher-Ilan (2001). Livnat's lessons, *Jerusalem Post*, June 19.
- Freeland, J. (1996). The Global, the National and the Local: Forces in the Development of Education for Indigenous Peoples - The Case of Peru. *Compare*, 26, pp.167-195.
- Golan-Agnon, D. (2005). Inequality in education and discrimination against Palestinian Arab students in Israel. In D. Champagne and Abu-Saad, I. (eds.) *Indigenous and Minority Education: International Perspectives on Empowerment*. Beer Sheva: Negev Center for Regional Development, Ben-Gurion University of the Negev, pp.201-214.
- Hadawi, S. (1991). *Bitter Harvest: A Modern History of Palestine* (4th edition), New York: Olive Branch Press.
- Haider, A. (2005). (ed). *Arab Society in Israel: Population, Society, Economy (1)* Tel-Aviv: Hakibbutz Hameuchad. (in Hebrew).
- Lewin-Epstein, N. and Semyonov, M. (1993). *Arabs in the Israeli Economy*. Boulder, Colorado: Westview Press.
- Lev-Ari, S. (2003). Know thy neighbor: The study of Arabic, Arab culture and the Koran could improve life in Israel, *HaAretz*, Feb. 26.
- Luke, A. (1988). *Literacy, Textbooks, and Ideology*. London: Falmer Press.
- Lustick, I. (1980). *Arabs in the Jewish State: Israel's Control of a National Minority*. University of Texas Press: Austin.
- Makkawi, I. (2002). Role conflict and the dilemma of Palestinian teachers in Israel, *Comparative Education*, 38 (1), 39-52.
- Mar'i, S. (1985). The future of the Palestinian Arab education in Israel, *Journal of Palestine Studies*, 14, pp. 52-73.
- Mar'i, S. (1978). *Arab Education in Israel*. Syracuse University Press: Syracuse, New York.
- Masalha, N. (1997). *A Land without a People: Israel, Transfer and the Palestinians*. London: Faber and Faber.
- Meehan, M. (1999). Israeli textbooks and children's literature promote racism and hatred toward Palestinians and Arabs, *Washington Report on Middle East affairs*, September, 19-20.

- Ministry of Education and Culture, (2004). *100 Basic Concepts*, Jerusalem: Ministry of Education and Culture. [Online] http://cms.education.gov.il/EducationCMS/Units/Mazkirut_Pedagogit/Mate/TochnitHaliba/TochnitLiba/Tokhnit100Musagim.htm [18 March 2006].
- Orni, E., and Efrat, E. (1992). *Geography in Israel*. Tel Aviv: Achiasaf. (in Hebrew).
- Peres, Y. Ehrlich, A. and Yuval-Davis, N. (1970). National education for Arab youth in Israel: A comparison of curricula, *Race*, 12 (1), 26-36.
- Podeh, E. (2002). *The Arab-Israeli Conflict in Israeli History Textbooks, 1948-2000*. Westport, CT: Bergin and Garvery.
- Raz-Krakotzkin, A. (1999). The Textbook Debate: No significant Change in the Perception of the History of Zionism, *Within Israel and Palestine*, XV (11), 1-6.
- Said, E. (1978). *Orientalism*. New York: Vintage Books.
- Sorkis, H., Raf, A., and Sharar, T. (eds.). (1991). *Changes in the Geography of Israel*. Jerusalem: Maalot. (in Hebrew).
- Swirski, S. (1999). *Politics and Education in Israel: Comparisons with the United States*. New York: Falmer Press.
- Zeidani, S. (2005). The Palestinian Arab predicament in Israel. In Y. Reiter (ed.) *Dilemmas in Arab-Jewish Relations in Israel*. Tel Aviv: Schocken Publishing House, pp.89-96 (in Hebrew).

Insisting on equity: A redistribution approach to education

Lee E. Nordstrum

Independent researcher, France lnordstrum@yahoo.com

Emphasis on educational efficiency, or accountability, with all its claims towards institutional improvement, has left the question of educational equity unanswered. Contemporary policy, particularly in Western nations, focuses largely on raising aggregate test scores while the greater society sees a steady increase in economic inequality. It is the purpose of this article to trace the current increase in inequality through income and wealth distributions, and argue that educational equity should also be a priority for all nations. Part one of this article gives both quantitative and qualitative examples of growing economic inequalities, and posits reasons why this is detrimental to human society. Part two provides evidence that an equitable redistribution approach to education may serve as part of an answer to these problems. Part three addresses counterarguments to these assertions, and describes how education alone cannot solve the question of equity.

Educational equity, funding, wealth and income inequality, economics, redistribution

INTRODUCTION

Working Toward a Definition of Equitable Education

In order to identify what is meant by 'equitable education', it is first necessary to distinguish it from a more common phrase: 'equal education.' Indeed, the two terms are often used interchangeably, as their technical definitions are similar, yet subtle distinctions arise in their implications. Equality, as it applies to education, implies sameness in all aspects; that is, it implies sameness in funding, curriculum, participation, and any other measure one could think of. Equal education means the same education for all.

For schools within a society to be equal, they must presently have (and always have had in the past) equality in three measurements of education: inputs, outcomes, and participation (Gylfason and Zoega, 2003). In other words, an equal state of affairs in education requires pre-existing conditions of equal funding, the same resources (monetary and technological), the same quality of teachers, the same quality instruction, the same number of years of instruction, and the same social outcome (social chance) of education.

It is clear that education is not the same (or equal) in all of these areas in any given society; there exists no nation that can lay claim to a completely equal distribution of education and human capital. Thus, there are some individuals who are at a disadvantage, who are on the losing end of this unequal distribution. In the United States, for example, a school district that pays its teachers \$77,000 per year will attract better and more educated employees than a district that pays \$36,000. A district that provides its children with a yearly education worth \$22,000 is bound to have greater access to meaningful technology and more resources than a district that provides an education worth only \$8,000 (Kozol, 2005). It is not enough that each child has access to free schooling.

Education in many developed nations is equal in that all students have the opportunity to attend some school through a certain amount of years, but attending school does not ensure an equal education. Moreover, if a district cannot even afford sanitary conditions for its students, how can it be expected to deliver the same instruction as a new, state of the art, resource-rich facility? How can students begin to take in this instruction, when their most basic needs (safety and health) are not being met?

Here, then, at the place where equality in education has failed and inevitably left behind those who are poor and who cannot gain access to quality schooling, is the case for equity. Equity carries with it a sense of justice and fairness as opposed to sameness. It takes into consideration that past circumstances may not have been equal, and implies just action is needed to restore fairness into the situation. It is equity, not equality, that suggests targeting poorer schools and districts with extra funds and resources may be necessary to “even the playing field.”

Equity speaks for honesty and integrity. It speaks to a system that needs to target those who start at a disadvantage through equitable redistributive policies. Equity is not equality, but that is its goal.

Equity as a Human Right

The assertion of educational equity as a human right is neither new nor original, but it is essential in formulating educational policy and has great implications, as this paper demonstrates. In a 2004 working paper, Kannan and Pillai describe three distinct generations of human rights (Kannan and Pillai, 2004). The first generation constitutes pre-twentieth century civil (or political) rights. Social and economic additions, such as the right to medical care, employment, food and shelter, came subsequently. The third generation of rights, ‘cultural rights’, includes the preservation of language, cultural institutions, and practices. Uniquely, the second generation rights are often portrayed as ‘goals’, rather than rights. Thus they are separate, not immediately binding, and only progressively realised. In fact, the European Convention on Human Rights of 1950 excluded them altogether, and the United Nations opted to include them in a separate treaty instead of the *Universal Declaration of Human Rights*¹. Naturally, this led to objections of granting welfare rights the status of human rights (Kannan and Pillai, 2004).

Kannan and Pillai rightly argued that welfare rights, such as an adequate standard of living, primary health care and public education, were closely related to the right to life. Individuals could not be free unless they were free from illness, have adequate access to food for survival and education for mental development. This view of welfare rights as civil rights has become increasingly dominant in the current landscape of literature (Sunstein, 2001). Nobel Prize winner (1998) Amartya Sen advanced this argument by distinguishing between availability and accessibility of social welfare. To Sen, goods, services, information, value, justice, recognition, and respect were not truly available unless they were made accessible through adequate purchasing power. Thus, individuals were free only to the extent that they had an unlimited “capability set” (1992). In this light, income deprivation is seen as a capability deprivation, which in turn, is a human rights violation.

If the definition of human rights includes welfare rights, and welfare rights include education, it follows that educational inequality of any kind is a human rights violation. Education begets an informed, rational choice of capabilities, which goes a long way towards determining freedom and well being (Kannan and Pillai, 2004; Sen, 1992, 1999). What was more, universal human rights, which included welfare rights, were not subject to popular or universal consent. If this were so, human rights might cease to exist altogether (Sen, 1999). One needs only to consider a

¹ Please see the *International Covenant on Economic, Social and Cultural Rights*, adopted and signed by the United Nations General Assembly on 16 December, 1966, and enforced on 3 January, 1976.

few international treaties to realise the universality of human rights is now taken as a given in modern society².

With education defined as a right to welfare, welfare rights defined as human rights, and human rights clearly defined as “universal, indivisible, interdependent and interrelated” (UNESCO, 1994), this article now turns to the problems posed by growing inequalities in the economic distributions of modern societies.

PART I – GROWING ECONOMIC INEQUITIES

Historical Background

Until quite recently, common economic thought held that, to a certain extent, prevalent inequality in a market system was good for economic growth. This was due in no small part to Kaldor’s historically significant hypothesis (1956): namely, that the marginal propensity of the rich to save and invest was greater than that of society’s poor. In this theory, an economy’s GDP was directly linked to the amount saved and invested. Therefore, it would be beneficial to instil mechanisms that inhibited income equality; it would be beneficial, economically speaking, to have an unequal distribution of wealth and income. A concentrated distribution would also ensure a greater amount of new industry and project development, as these endeavours required large sunk costs that wouldn’t be available to a more equitable distribution. Furthermore, from an incentives point of view, a concentrated distribution was good in that it showed the worker that output or success depended largely on observable input (Mirrlees, 1971). In this scheme, individuals had greater incentives to work hard.

Such thought was bolstered and moralised by Kuznet’s hypothesis, put forth in a famous 1955 article³, which stated that inequality was an essential and unavoidable result of economic growth, but would eventually peter out. Thus, a natural so-called ‘trickle-down’ effect would be seen in income and wealth inequalities in developed nations, justifying the earlier stage of inequality. Since there was a natural economic trend for inequality to fall, there was little need for government assistance or aid. This hypothesis largely explained the decrease in inequality from the 1930’s to the 1970’s in the United States to levels lower than both Great Britain and Sweden (Aghion, Caroli, and Garcia-Penalosa, 1999).

Current Trends

Though Kuznets’ hypothesis was essentially granted the status of economic law, it did not have sufficient evidence to support its claims, (HRSDC, 2000), nor did it explain the rise of inequality in several developed nations since the 1970s. Since that time, income and wealth inequality has grown to such an extent that the top five per cent of the income bracket own 59 per cent of all the wealth in the United States (Wolff, 2003). In other words, the top five per cent own more wealth than the rest of the United States does collectively. Measured by the GINI coefficient⁴, in which 100 is the highest possible level of inequality, this wealth distribution registers 86, almost

² Please see the *Universal Declaration of Human Rights* (United Nations, 1948), specifically Articles 1 and 2, and the UNESCO Declaration of the 44th Assembly (UNESCO, 1994).

³ Kuznets’ article was entitled “Economic Growth and Income Inequality” and was printed in the *American Economic Review* (1955).

⁴ The GINI coefficient is a calculation of the distribution of income within a society on a scale ranging from 0.00 (perfect equality) to 100.0 (perfect inequality). Specifically, it calculates the extent to which an economy’s income distribution (graphed as a ‘Lorenz Curve’) deviates from a hypothetical line of perfect equality, and is expressed as a percentage of the maximum area under that line (UNDP, 2003). GINI coefficients can be calculated for both income and wealth distributions. In 2003, the United States recorded a GINI (income) of 40.8, significantly above the top five most equally distributed countries (all which have GINI indexes hovering around 25.0), and tying its income distribution with that of Turkmenistan (US Bureau of the Census).

perfectly unequal. When the data are broken down by race, Wolff finds that the “average African-American family has about 60 per cent of the income of the average white family. But the disparity of wealth is a lot greater. The average African-American family has only 18 per cent of the wealth of the average white family” (Wolff, 2003).

Perhaps more disturbing is the persistence of child poverty in developed nations despite the doubling and redoubling of per capita incomes since the middle of the twentieth century (UNESCO, 2000). This finding, based on conditions in the world’s richest countries, is particularly troubling to traditional economic thought, for if Kuznets is correct, and inequality in fact falls with the rise of per capita income, the problem of poverty should not be as pervasive as it is⁵. In fact, the report found that 47 million children, or one in every six, from OECD countries were living below their country’s poverty line (UNICEF, 2001). In the study’s own words:

The persistence of child poverty in rich countries undermines both equality of opportunity and commonality of values. It therefore confronts the industrialised world with a test both of its ideals and of its capacity to resolve many of its most intractable social problems.

According to the study, this persistence was not due to the ongoing debate surrounding differing measures of poverty. To be true, no matter which measure was used, either absolute or relative, the top six best performing countries remained the same: Sweden, Norway, Finland, Belgium, Luxembourg, and Denmark. These countries were among the world’s wealthiest, and yet they had found a way to combine a high degree of equity and economic development. In other words, policy could have an affect on inequality and child poverty did not have to persist in the numbers that it did.

Traditional economic theory would perhaps not deem this a problem, as inequality (even instances of poverty) is considered instrumental for higher economic growth patterns. However, this conclusion is hindered by several implications, based both on recent empirical data and popular sentiment. The first is a moral implication: many people object to such high levels of inequality and access to well being within their society (Wolff, 2003). Green, Preston and Sabates (2003) affirm that persistent inequality at the present time is one of the factors contributing to societal tension and lack of cohesion.

Second, unequal economic distributions correlate to other unequal distributions in resources, educational achievement, and benefits. More equal societies tend to have more equal distributions of human capital, leading to a much more educated workforce. Unequal societies have just the opposite: a workforce that is not as well prepared and consequently not as productive (Wolff, 2003). Wolff also states that aggregate educational achievement suffers in a nation with grand disparities in educational access, as opposed to more equitable societies (2003). To this end, a recent study highlights the existing educational gap between English Language Learners and their English speaking peers in California. The vast majority of English Language Learners, who, in California’s case, tend to be immigrants from Mexico and significantly poorer, are more likely to receive fewer educational resources and achieve lower results than English speaking children (Gandara, Rumberger, Maxwell-Jolly, and Callahan, 2003). According to the study, this gap

⁵ This is based on a UNICEF study entitled *Child Poverty in Rich Nations*, which uses the measure of relative poverty as opposed to absolute poverty. Debate exists on which measure is best, and the two give differing answers as to the pervasiveness of poverty, but UNICEF defends its decision for several reasons. First, relative poverty is the most widely used measure of poverty, and thus provides the best opportunity for comparison; only the United States uses a measure of absolute poverty, which is then converted to differing currencies and incomes. Second, the definition of absolute poverty was developed in the 1960’s based on the dollar amount of an adequate diet, multiplied by three. However, changes in the cost of living have resulted in the cost of an adequate diet no longer constituting one-third of an adequate income. Third, the effect of poverty in its absolute form is not the main concern; instead, it is the daily-perceived contrast of the way of lives of a country’s poor.

manifests itself as unqualified teachers for English Language learners, an inferior curriculum with less time to cover it, inferior facilities, invalid assessment instruments, and segregation from their peers.

The third implication of a concentrated income and wealth distribution is hampering economic performance and growth. A growing body of respected research literature questions Kaldor's hypothesis, and depicts inequality as negative for economic growth (Aghion et al., 1999; Alesina and Rodrick, 1994; Perotti, 1993, 1996; Persson and Tabellini, 1994). Together, these empirical studies site two potentially positive effects of redistribution in a market economy. The first, an '**opportunity enhancing effect**' states the lowest investment of the rich is not productive and gives decreasing returns. The investment productivity made by the poor, on the other hand, is relatively high, yet they are unable to invest beyond their limited endowments. Thus, if the lowest investments of the rich were redistributed to the poor, productivity would increase, creating an obvious benefit for the poor, an insignificant change in wealth for the rich, and a net benefit for the economy (Aghion et al., 1999). The second, a '**positive incentive effect**', claims that high levels of inequality produce inefficiencies, in that there is an inverse relationship between the number of individuals with low incomes and the amount of effort put forth in the workplace (Aghion and Bolton, 1997; Banerjee and Newman, 1993). Again, redistribution does not negatively influence the production of the rich so much as it positively influences the production of the poor, resulting in a net gain.

Finally, it is important to note that all of the studies cited in this section, excluding Kaldor and Kuznets, favoured a redistribution of resources from the wealthiest individuals to the poor. Such a policy proves a beneficial force for both the individual and the greater society in which it rests, by freeing the capability set of the poor, promoting access to fundamental resources, and conditioning greater economic growth. In section two, this paper describes how this can be done through education.

PART II – EDUCATION AS AN ANSWER TO ECONOMIC INEQUITIES

Despite the ingrained traditional economic thought, this article points to numerous studies that cite new patterns in growth and inequity, and link more equitable economies with higher rates of growth. Barro (2000) found that inequality retarded growth in poor countries, a finding directly opposed to the traditional hypotheses. Garcia-Penalosa (1995), too, found that inequality in rich nations reduced growth, as it increased the number of poor who could afford neither to invest nor to save. More studies demonstrated the growth hampering effects of income and wealth inequalities in light of market imperfections, thus making a case for redistribution (Aghion, Caroli, and Garcia-Penalosa, 1999; Galor and Ziera, 1993).

Recently, Gylfason and Zoega (2003) has extended the assertions above to test the relationship between human capital, or education, and economic growth. By comparing cross country data on GINI coefficients, GDP, GNP, government educational expenditures, and length of basic education, they are able to observe a similar relationship between human capital and growth. Though other authors discount the potential role education plays in growth analysis by claiming diminishing returns (Aghion et al., 1999), Gylfason and Zoega emphasise the 'dual effect' of education expenditures (2003). Consider, for a moment, a redistributive policy of human capital in a society that has a concentrated wealth and educational distribution, taking monetary resources from the elite and giving them to the less educated. Such a 'Robin Hood' approach is justified, as spending monetary resources on those less educated generates more human capital than if that same amount is spent on the education of the elite. It follows, then, that the improvement of schools providing basic education, at the expense of elite funding, enhances both equality and human capital. Now the question arises: what does this scheme do to growth? Gylfason and Zoega see two growth inducing effects (2003). In the first, education's increase in

human capital leads to an increase of physical capital and gives the nation a greater net propensity to save and invest. In the second, this increase in human and physical capital, along with the heightened educational level of the country, produces more requests for progressively equitable conditions; a virtuous circle and greater demand for equity is created, which in turn increases growth even further.

The previous claims flow from patterns Gylfason and Zoega observe in cross national data, which seem to support the view that inequality goes with less investment and smaller growth. When comparing GINI indexes, GDP and GNP, a rise in a country's GINI coefficient by 12 points correlates with a drop in both GDP and GNP by 1.5 per cent and 1.0 per cent, respectively. The ensuing two findings, however, are perhaps the most significant for this paper, as they directly implicate a correlation between education, growth, and inequality. Gylfason and Zoega find that an increase in educational expenditures, expressed as a percentage of GNP, correlate with a drop in GINI coefficient and a rise in economic growth. Specifically, a rise of one per cent in GNP in educational spending correlates with a drop of 2.3 points on the GINI scale, and a three per cent rise in GNP spending correlates with an economic growth increase of one per cent. Not surprisingly, the authors also find that an increase in years of boys and girls' schooling correlates with a GINI decrease and increase in growth. What is surprising, though, is that as little as a one year increase in girls' education correlates with a decrease in the GINI coefficient and an increase in economic output.

Two conclusions can be drawn from these data, both of which are important in terms of educational and economic policy. First, economic growth varies inversely with inequality. Second, of three measures in education (input, outcomes, and participation), all are inversely related with inequality, and economic growth varies directly with these three indicators. While recognising that correlations do not necessarily imply causation, these findings give credit to the assertion that efforts to raise aggregate and individual levels of human capital have an impact on the societal level (Gylfason and Zoega, 2003). Indeed, education seems likely to encourage economic growth by increasing and improving human and physical capital.

Put together, Gylfason and Zoega's findings advocate a redistribution approach to education. The equitable distribution of educational spending is an underlying assumption in their conclusions; without this their correlations fall away. To use an extreme illustration, a country that increases its GNP educational expenditures, but gives 95 per cent of it to half of its school districts and only five per cent to the remaining half is not likely to see a corresponding decrease in GINI coefficient. More realistically, imagine a country that increases its GNP expenditure and gives an equal monetary amount to each school district. Such a country inspection to see a decrease in the GINI coefficient and a rise in economic growth. However, if that country has districts that are wealthier than others, the ensuing situation is the same; there is an inflated resource level across all districts, but a funding gap remains. The only possible way to beget a decrease in GINI levels and an increase in economic growth is to increase GNP educational expenditures in accordance with an equitable redistribution approach that targets districts behind the funding gap.

In order to raise levels of human capital to a more equitable level and potentially impact economic growth, education must be equitable in all three measures proposed above. Inputs, outcomes, and participation must be equitable, or human capital continues to be a concentrated distribution. If efforts to raise human capital are unequal, if there are any community pockets left untouched by educational reform, it is unavoidable that the ensuing human capital is not equitable. Such a situation does nothing to alleviate the current one. In order for positive effects to be seen, educational reformers must insist on a redistribution scheme as outlined above.

PART III – COUNTERARGUMENTS AND SYNTHESIS

Counter arguments to this paper's position run along a few distinct lines: (a) the recent growth of inequality and particularly poverty; (b) the assertion that inequality is bad for economic growth; and (c) the increase of education (or human capital) as influential to economic growth. The first contention, the increase of poverty and inequality, is often a case of differing measurements. The United States uses an absolute measure of poverty, wherein poverty is defined as a family living below a certain set income, whereas the rest of the world, UNICEF, UNESCO, and the World Bank use a measure of relative poverty. Here, poverty is defined as a family living with less than 50 to 60 per cent of the nation's average income. The two measures provide dissimilar numbers of poverty stricken families and individuals in a given country and paint different pictures of societal inequality. The United States, in particular, suffers a true fall from grace in the switch from an absolute to a relative measure of poverty (UNICEF, 2001). Relative poverty is important because most often it is not the physicality of starvation that affects the poor, but the relative chasm of difference that separates them from their peers, and does not allow them to live within their society's modern framework of the so-called 'good life' (UNICEF, 2001). The stigma of poverty stays with them.

Many modern studies have sighted income and wealth inequalities as bad for economic growth. However, others argued against this conclusion, such as Forbes (2000), who found a positive relationship between inequality and economic growth, forming the second counterargument. Forbes' findings, would be thrown into question by the findings of an earlier study (Aghion et al., 1999). These authors would challenge her analysis on the basis that her sample population was excessively small, and was of greater significance. In addition, they challenge that there was little evidence behind Forbes' so-called 'lag structure,' or her assumption that inequality today affects economic growth in five years. Last, in order to obtain a positive coefficient between inequality and growth, Forbes restricted data to countries listed in the 'high quality' subset posited by Deininger and Squire (1996). Aghion et al., (1999) on the other hand, showed the criterion used to construct this subset was inappropriate, and that there was no reason to exclude other countries.

Most significantly, the study that this article so heavily draws on in the first part (Aghion et al., 1999), argued that increasing human capital did not necessarily work to decrease inequality. To the authors, educational investment yielded decreasing returns, and its effects on inequality differed greatly depending on the society and the system of education. For example, the United Kingdom and the United States reacted to more human capital by firing the less skilled workers and hiring those with greater skills. Japan and Germany, however, emphasised experience and tended to promote from within. In these cases, an increase in human capital might increase job skills and preparedness, but nations then turned to a different emphasis in the hiring process, namely, 'experience'. Equitable education and an increase in human capital might cause an emphatic shift from educational attainment to experience, but there was still a similar amount of inequality. There was a different cause with the same net effect.

CONCLUSIONS AND RECOMMENDATIONS

It is not intended, by addressing counterarguments to this article's position, to somehow weaken the assertion that education is an essential part of addressing societal inequities. Instead, the arguments are useful to highlight another aspect relevant to this topic: though education may be a tool, it is not the only tool, and in some cases certainly not the sharpest, to cut across the woven fabric of inequality within societies. The correlations between education, economic growth, and income and wealth distributions do not prove that education is the root cause of all positive effects in these categories, but they do show that equitable education is associated with them. In considering social well being and the potential benefits it brings to a nation, the most important question is not whether education or income equality is a stronger factor. The most important

conclusion is that **both education and income equality** may have an effect on societal well being. Education plays a part in the alleviation of societal maladies. It is not the only part by any means, but it is part of the web of correlations that could provide a more egalitarian, more efficient, more cohesive society for future generations. Current trends are pushing the other way: there is less cohesion (Green et al., 2003) and more inequality (Aghion et al., 1999; Gylfason and Zoega, 2003; Wolff, 2003). These trends do not exist in a vacuum, rather, they are interrelated. So, too, educational policy is not meant to exist in a vacuum. The best policy, then, is then to synthesise findings in educational, social, and economic research and cooperatively to combat inequities.

In conclusion, it is essential to remember that education must be equitable in three aspects: inputs, outcomes, and participation. If it fails to be equitable in any of these, its potential positive returns are effectively annulled. Equity is the key. Equity addresses inequalities in its educational outcomes through redistribution, provides more equal opportunities to an egalitarian income and wealth distribution, is good for economic growth, and increases societal cohesion (Green et al., 2003).

Raising aggregate test scores and average educational achievement do not by themselves have these effects; raising mean scores do not address inequities that plague today's societies. If scores are raised without an ethos of equity and redistribution, achievement gaps between classes and races will still persist (Gandara et al., 2003). Politicians and leaders can cite statistics that illustrate that scores are rising, but if they do not address inequities, their people will not listen. It does not matter what statistics say, because the more influential stigma of relative inequality remains as a distasteful moral problem for some, and a pervasive glass ceiling for others. Statistics, whether they are rising or falling, whether they are deemed good signs or bad, only matter in so far as they affect the lives of individual citizens. Therefore, high standards and high aggregate test scores must be sought after **in conjunction with the eradication of inequities through education expenditure redistribution.**

This article ends by providing several recommendations, based on evidence described above, for future education policy and research.

1. Nations, their citizens, academics, and policy makers must insist on educational equity in three aspects (inputs, outcomes, and participation), not just raising mean levels of achievement. This can and should be done while keeping standards for educational instruction and curriculum high.
2. If equity is seen as a 'goal' in education and is not present from the beginning, it is much more difficult to attain later. As noted, there are many contemporary forces pushing toward rising inequities. If equity is not present at the beginning, it is seen as peripheral and not essential. Thus, inequity slowly but surely becomes accepted by society's constituents as the *status quo*, effectively turning a nation into a nation of exclusion.
3. Measures of educational equity must be improved. The United States boasts a GINI educational index of 15.8, which is relatively close to perfect equality. However, educational outcomes are by no means equal across the country, mainly because of funding inequities. The upshot is that students from wealthy backgrounds have a much superior education than students from poor backgrounds. Future educational equity measures must take this into account.
4. This paper encourages more research into the potential benefits of redistribution in human and physical capital within nations (rich and poor). It may in fact turn out that such policies benefit both aggregate knowledge and educational equity. Policy, then, must be reformed to fit these findings. Here, the studies of Scandinavian economies have a great deal of knowledge to offer.

5. Last, this paper encourages a continuing dialogue on the importance of equity, at both the academic and societal levels. More than lip service, this dialogue is a vital part of a democratic society and must be taken seriously by everyone, from citizens to policy makers. Inequity affects whole societies sociologically, economically, politically and philosophically, but its effects are most palpable and devastating to the individual. It is time they are given adequate voice.

REFERENCES

- Aghion, P. and Bolton, P. (1997). A trickle-down theory of growth and development. *Review of Economic Studies*, 64, 151-172.
- Aghion, P., Caroli, E., and Garcia-Penalosa, C. (1999). Inequality and economic growth: the perspective of the new growth theories. *Journal of Economic Literature*, 37 (December), 1615-1660.
- Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109 (2), 465-490.
- Banerjee, A. and Newman, A. (1993). Occupational choice and the process of development. *Journal of Political Economy*, 58, 211-235.
- Barro, R.J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5 (1), 87-120.
- Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *World Bank Economic Review*, 10 (3), 565-591.
- Forbes, K.J. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review*, 90 (4), 869-887.
- Gandara, P., Rumberger, R., Maxwell-Jolly, J., Callahan, R. (2003). English learners in California schools: unequal resources, unequal outcomes. *Education Policy Analysis Archives*, 11 (36). [Online] http://www.usc.edu/dept/education/CMMR/FullText/ELLs_in_California_Schools.pdf [15 November, 2005].
- Garcia-Penalosa, C. (1995). The paradox of education or the good side of inequality. *Oxford Economic Papers*, 47, 265-285.
- Galor, O., and Ziera, J. (1993). Income distribution and macroeconomics. *Review of Economic Studies*, 60, 35-52.
- Green, A., Preston, J., and Sabates, R. (2003). Education, equality, and social cohesion: a distributional approach. *Compare*, 33 (4), 453-470.
- Gylfason, T., and Zoega, G. (2003). Education, social equality, and economic growth: a view of the landscape. *CESifo Economic Studies*, 49 (4), 557-579.
- HRSDC. (2000). The impacts of inequality on productivity growth: a primer. [Online] <http://www11.hrsdc.gc.ca/en/cs/sp/hrsd/arb/publications/research/2000-000022/page00.shtml> [10 January, 2006].
- Kaldor, N. (1956). Alternative theories of distribution. *Review of Economic Studies*, 23, 83-100.
- Kannan, K.P., and Pillai, V. (2004). Development as a right to freedom: an interpretation of the 'Kerala Model.' [Online] http://www.cds.edu/download_files/361.pdf [21 October, 2005]
- Kozol, J. (2005). Still separate, still unequal. *Harper's Magazine*, 311 (1864), (1 September, 2005).
- Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45, 1-28.
- Mirrlees, J.A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38 (114), 175-208.
- Perotti, R. (1993). Political equilibrium, income distribution and growth. *Review of Economic Studies*, 60, 151-172.
- Perotti, R. (1996). Democracy, income distribution and growth: what the data say. *Journal of Economic Growth*, 1 (June), 149-187.

- Persson, T. and Tabellini, G. (1994). Is inequality harmful for growth? *American Economic Review*, 84 (3), 600-621.
- Sen, A. (1992). *Inequality Re-examined*. New York: Oxford University Press.
- Sen, A. (1999). Democracy as a universal value. *Journal of Democracy*, 10 (3), 3-17. [Online] <http://muse.jhu.edu/demo/jod/10.3sen.html> [14 November, 2005]
- Sunstein, C. (2001). *Designing Democracy*. New York: Oxford University Press.
- UNESCO (1994). Declaration and Integrated Framework of Action on Education for Peace, Human Rights and Democracy. [Online] http://www.unesco.org/human_rights/hrff.htm [Accessed 9 December, 2005]
- UNICEF (2000). A league table of child poverty in rich nations. *Innocenti Report Card*, 1, (June 2000).
- United Nations (1948). Universal Declaration of Human Rights. (Adopted 10 December, 1948). [Online] <http://www.unhchr.ch/udhr/index.htm> [9 December, 2005]
- United Nations (1966). International Convention on Economic, Social, and Cultural Rights. (Adopted 16 December, 1966). [Online] <http://www.ohchr.org/english/law/cescr.htm> [9 December, 2005]
- Wolff, E. (2003). The wealth divide: the growing gap in the United States between the rich and the rest. *Multinational Monitor*, 24 (5). [Online] <http://multinationalmonitor.org/mm2003/03may/may03interviewswolff.html> [10 December, 2005]

Retention and academic achievement research revisited from a United States perspective

Jon Lorence

Department of Sociology, University of Houston jlorencel@uh.edu.usa

Educational researchers in the United States contend that making low-performing students repeat a grade is an ineffective educational practice. This view derives largely from the summary of grade retention research reported by Holmes (1989). A meta-analysis of more recent studies (Jimerson, 2001) also concludes that the practice of grade retention should be abandoned. However, a thorough examination of the published articles within each of these two meta-analyses reveals that many of the individual studies evidence inadequate research designs and faulty conclusions. The overwhelming majority of conclusions from grade retention studies are unwarranted due to the poor quality of research. Overlooked and more recent retention and grade repartition studies suggest that making students repeat a grade may help increase academic achievement. This review contends that research studies do not support the contention that grade retention is always inappropriate. Suggestions for improving future retention studies are offered.

Grade retention, academic achievement, meta-analysis, faulty conclusions, inadequate research designs, grade repeating

BACKGROUND

The overwhelming majority of educational researchers have concluded that requiring low-achieving students to repeat a grade is an inappropriate, if not harmful, educational practice. Mantzicopoulos and Morrison (1992) contend that “Unlike mixed empirical evidence on other educational issues, research on elementary school nonpromotion is unequivocal. It supports the conclusion that retention is not an effective policy” (p.183). Two of the most vocal opponents of grade retention practices argue that “retention worsens rather than improves the level of student achievement in years following the repeat year” (Shepard and Smith, 1990, p.88). Besides dominating educational research journals, this highly critical view towards making academically-challenged students repeat a grade pervades publications which address the practical concerns of teachers and educational administrators (for example, see Darling-Hammond and Falk, 1997; Harrington-Lueker, 1998; Owings and Magliaro, 1998; Potter, 1996; Reynolds, Temple, and McCoy, 1997). The National Association of School Psychologists view the practice of grade retention to be so ineffective that the organization “urges schools and parents to seek alternatives to retention that more effectively address the specific instructional needs of academic underachievers” (NASP, 2003).

Given such strong beliefs, one would assume that the research demonstrating the futility of grade retention or grade repetition would be equally compelling. However, there is some disagreement regarding the persuasiveness of conclusions commonly associated with grade retention studies. In the most comprehensive review of its time examining the impact of making low-performing students repeat a grade, as opposed to being promoted to the next grade, Jackson (1975, p.627) concluded “that the accumulated research evidence is so poor that valid inferences cannot be drawn concerning the relative benefits of these two options.” More recent meta-analyses of grade

retention or grade repetition research (for example, Holmes, 1989; Jimerson, 2001), however, assert that the practice of requiring students to repeat a grade is ineffective at best and likely to be detrimental to further student academic achievement. Conversely, Lorence, Dworkin, Toenjes, and Hill (2002) offered a cursory critique of the grade retention literature and questioned the conclusions of various studies opposed to holding students back in grade for an additional year. Alexander, Entwisle, and Dauber (1994) have been even more critical of the grade retention literature; they argue that an overwhelming number of retention studies are flawed, resulting in erroneous interpretations based on “bad science” (p.220). Shepard (2002, p.57) disagreed with these critical assessment of grade retention studies and accused those questioning the negative impact of grade retention for ignoring the “weight of evidence” which purportedly demonstrates the failure of retention practices. Because educational researchers are likely to be more familiar with the negative conclusions of the grade retention summaries, rather than the specific studies which compose the Holmes and Jimerson meta-analyses, it is worthwhile to explore in greater detail the retention literature. Consequently, readers should not assume that the purpose of this review is to argue for retaining academically challenged students. The major intent of this review is to make readers aware that most of the studies cited in the retention literature are insufficiently sound to support the contention that making students repeat a grade is always wrong.

Although Alexander et al. (2003, chap. 2) point out general shortcomings of the grade retention literature, the current paper presents a more thorough examination of the many specific articles that attempt to assess the effect of retention on student learning outcomes. The purpose of the current paper is to examine the foundation upon which is based the prevailing assertion that requiring students to repeat a grade will not contribute to their academic progress. First specified are criteria used to evaluate the quality of grade retention studies. In addition to usual issues related to the quality of research, a discussion of the appropriate strategy to compare the academic progress of retained and nonretained students is presented. The Holmes (1989) meta-analysis is then scrutinised because his article is often cited as the definitive study which demonstrates that requiring students to repeat a grade is an ineffective educational practice. Next examined is Jimerson’s (2001) review which similarly concludes that grade retention policies should be abandoned. An examination of overlooked and new studies pertaining to retention then follows.

CRITERIA OF RESEARCH QUALITY

Published Research

Dunkin (1996) identified nine common types of errors which can occur when attempting to synthesise educational studies (see also Wolf, 1986). Although not all of the problems which can arise in meta-analyses are mentioned here, several errors are pervasive throughout the grade retention literature. The most problematic feature of the retention summaries is that the quality of the individual studies varies considerably; oftentimes mediocre studies are weighted equally with good research. A commonly used indicator of research quality is the type of outlet in which a study appears. Alexander et al. (1994) as well as Dunkin (1996) state that published findings which have undergone a review process are more likely to be of higher quality than unpublished convention papers, master’s theses, or dissertations.

Although many readers would likely agree that published findings are probably of higher quality than nonreviewed papers, less consensus exists with respect to other criteria of research quality. Dunkin (1996) warns that assessing the quality of a study “is possibly the most difficult type of error to identify, because criteria for evaluating research are usually controversial, and judges sometimes disagree on the difference between good and poor research” (p.89). To illustrate, critics of grade retention practices assume that the research designs of most studies are appropriate and yield valid results. For example, Heubert and Hauser (1999, p.122) and Shepard (2000) contend that the Holmes (1989) meta-analysis clearly supports the notion that grade

retention is harmful to students because the sum of all the “effects” across the many studies examined are negative. Jimerson (2001) reaches a similar conclusion by counting the number of positive and negative effects calculated from the retention studies he reviewed. His “scorecard” approach yields a greater frequency of negative effects between retained and promoted students. Moreover, average weighted effect sizes Jimerson (2001) presents are also negative. Grade retention is therefore presumed to be an inappropriate educational practice. However, readers seldom question the quality of studies composing these meta-analyses. In addition to the type of publication outlet as an indicator of quality, additional criteria offered to gauge the value of grade retention findings are; research design, type of statistical analysis, comparison strategy, scale of measurement, sample representativeness, and sample size.

RESEARCH DESIGN – MATCHING VERSUS STATISTICAL CONTROLS

Probably the most important criterion for judging the quality of a study is the nature of the research design utilised. Researchers would agree that the best way to assess the effect of different educational practices is by use of an experimental design in which low-performing students are randomly assigned to clearly delineated treatment conditions. Random assignment would help equalise initial differences among students prior to retention which, if ignored, might result in misleading conclusions regarding the effects of grade retention. Observed differences at the end of the study period between retained and promoted students might arise from initial dissimilarities in student abilities rather than retention practices. Although Jackson (1975, p.628) recommended the use of experimental models to assess the effectiveness of grade retention, ethical and practical considerations preclude the implementation of true randomised experimental methods. Most parents and teachers are unwilling to allow random assignment of pupils because of the fear that some students would be harmed if retained (or not retained). Further, random assignment would not be consistent with individual school district promotion policies and practices used to assign pupils to specific grades as well as teachers.

Such constraints have forced researchers to adopt research designs which attempt to approximate the contours of a true randomised experiment. All of the studies cited in the aforementioned meta-analyses, resemble the quasi-experimental “nonequivalent control group design” described by Campbell and Stanley (1963, pp.47-50). In all the retention studies reviewed in this paper, at least two groups of students were studied -- those who were required to repeat a grade and those who were promoted to the next grade. Differences in outcome measures were then examined at the end of the experimental period, usually one or more years after the time of retention.

Controlling by Matching

Due to the inability to randomly assign students to the promoted or retention condition, researchers attempt to equalise initial differences between the two sets of students by the common practice of matching. That is, promoted students with certain characteristics (for example, age, gender, race, and socio-economic standing) presumed to be similar to the retained pupils were identified in school records: Their outcome scores were then compared with those of the retained students. While the strategy of matching is often the only procedure available, most researchers acknowledge that conclusions from such studies should be interpreted cautiously. To illustrate, Campbell and Stanley (1963, p.15) argued that “matching is no real help when used to overcome initial group differences.” Campbell and Kenny (1999) later warned: “The danger of matching is that, although the scores are more equivalent due to matching, it is unlikely that they are exactly equivalent. Thus, matching achieves more the illusion of equivalence than the reality” (p.54). With respect to retention practices, Jackson (1975, p.619) pointed out that studies which compared retained students matched with promoted pupils are biased against those held back because initial differences between the two kinds of students will not be adequately controlled by

the practice of matching. Although students may be more similar if matched, promoted pupils will still not be comparable on the most important variables related to retention because those students held back will have greater difficulties in unspecified areas; otherwise they would not have been retained. Alexander et al. (2003) also contend that “It is impossible to match retained and promoted students on all relevant factors – the promoted group, for example, may be more mature, have fewer family problems, or be less aggressive than the held-back group” (p.25). If students are to be equated through matching, a measure of the outcome variable, or at least an indicator of student ability, should be the basis for determining comparability of retained and promoted pupils. In this paper the minimum requirement to be considered appropriately matched is that the two groups should be statistically similar on ability level and/or the initial indicators of the outcome measures. If retained and nonretained students are not comparable, then initial differences should be at least adjusted through statistical methods.

Controlling by Statistical Adjustment

Related to the issue of nonrandom research designs is the choice of appropriate statistical procedures to analyse findings. If one were certain that the retained and promoted students were truly equivalent prior to the year in which pupils repeated a grade, the post-treatment means of the two groups could be compared with a *t* test. Statistically significant differences between groups at the end of the study would then suggest the nature of the effect of retention. Insofar as it is highly unlikely that retained and promoted students are equal in performance at the time of retention, outcome comparisons between groups based on mean differences will be inaccurate and misleading. Campbell and Stanley (1963, p.49) therefore recommended that, when possible, an analysis of covariance be used to help statistically adjust for potential initial differences between nonequivalent groups. More recently Campbell and Kenny (1999, pp.68-79) demonstrated how the use of covariates as statistical controls can help overcome potential selection effects which may result in a better estimate of the impact of a treatment on outcome measures. While many variables could be used for statistical adjustments between groups, it is most important to have a pretreatment measure of the outcome variable because statistically controlling for this pretest will help ensure a truer gauge of the treatment effect, especially if the groups differed on the outcome measure prior to placement in a control or experimental conditions (Maxwell and Delaney, 1990, chap. 9). Other variables which may influence the outcome measure should also be included in the analysis as covariates. As will be seen, only a small number of retention studies focusing on academic achievement attempted to statistically control for initial differences between retained and promoted students, even when such baseline information was available.¹ Although statistical adjustments can yield a more accurate assessment of the effect of making students repeat a grade, traditional linear statistical models can only partially control for initial differences between retained and promoted pupils. The shortcomings of statistical controls will be discussed in more detail after a review of the available retention studies. In spite of problems associated with statistical controls, particularly the availability of variables indicating preexisting differences between retained and nonretained students, many studies cited in the meta-analyses did not adjust for preretention differences even when appropriate variables were available.

GROUP COMPARISON STRATEGIES

Retention studies differ with respect to the time at which retained and nonretained students are compared. Several strategies have been used to assess whether promoted or nonpromoted pupils

¹ Maxwell, O’Callaghan, and Delaney (1993) speculated that educational researchers’ extensive use of matching techniques during the 1970s and 1980s stemmed from a number of articles critical of covariance methods to statistically equalise initial differences between treatment and control groups. However, Maxwell and Delaney (1990, chap. 9) pointed out that some of the earlier criticisms were overly stringent and that analysis of covariance procedures can be useful in helping interpret results from nonrandomized research designs.

have advanced relative to one another. Holmes (1989) categorised retention studies as using either a “same-age” or a “same-grade” comparison procedure. The former procedure compares the outcome measure(s) between the retained and promoted students during the same academic year, however, the nonpromoted pupils will be one grade behind their promoted classmates at the time of comparison. Students will be the same age but in a different grade. For example, assume academically challenged first graders were required to repeat the grade. At the end of the repeated first-grade year (that is, the first graders had sat through the same grade twice), the means of retainees would be compared with the average outcome measures of their previous classmates who would be in second grade. Assuming most of the retained and promoted children were of the same age in the initial first grade, the retained students would be of the same age of the nonretained students who had been promoted to second grade. The term “same-year” comparison is also sometimes used to specify that outcome measures are being made between groups of students who are not in the same grade (Karweit, 1992).

Same-grade comparisons, as defined by Holmes (1989, p.21), entails examining the mean performance measures of retained and promoted students when they are in the same grade, but not in the same year. Continuing with the previous scenario, assume the retained first graders had now completed second grade and their initial classmates were one year ahead finishing third grade. Contrasting the second grade scores of the initially retained children with the previous year’s means of their promoted classmates, when they were in second grade, results in a same-grade comparison. The time order of the mean outcome measures for the retained students will usually be one year behind that of the promoted pupils in a same-grade comparison. Occasionally, however, the term “same-grade” comparison refers to examining the performance of retained students with their current classmates who had never been required to repeat a grade. The classmates of the retained pupils will usually be one year younger even though they are in the same grade.

The comparison strategy selected for evaluating the impact of grade retention may influence the substantive conclusion of the analysis. Holmes (1989, pp.21-22) summarised that same-age comparisons yielded findings indicating that grade retention results in a negative effect on student outcome measures. Conversely, same-grade comparisons initially support grade retention. However, the positive effect of making students repeat a grade is assumed to quickly diminish in a few years as the gap between those held back and the promoted pupils decreases. Alexander et al. (2003, pp.22-23) speculated that the initial higher scores of retained students observed in same-grade comparisons at the end of the retention year (for example, first grade means of promoted pupils with the second set of first grade means of the students who repeated first grade) occurs because retainees have taken the same curriculum twice. Alternately, promoted students may evidence higher scores than their retained counterparts when performing a same-age comparison because the nonretained students have covered more classroom material due to an additional year of school in a more advanced grade. The promoted students have had an additional academic year of new curriculum than their retained prior classmates.

Given that the basis of comparison may bias findings that support or do not favour retention, it is important to decide whether the same-age or same-grade comparative strategy is substantively more appropriate. Wilson (1990) contends same-grade comparisons are more reasonable because same-age comparisons require retained students to obtain scores comparable to the promoted students, who have been exposed to an additional year of new curriculum. Retained students are in effect required to learn the next year’s curriculum in order to obtain the achievement scores similar to the promoted pupils if a same-age comparison is used. Karweit (1999, pp.43-44) likewise argues that parents and teachers are more concerned with how retained children perform relative to their current classmates rather than their previous peers who were promoted to a different grade. If one assumes that the purpose of making students repeat a grade is to allow them to learn the material necessary to proceed successfully in later grades, then the same-grade

comparison is the more appropriate basis for assessing the usefulness of grade retention. Even Shepard (2004), a major critic of retention, agrees that “same-grade comparisons fit better the logic of retention, which is intended to be a one-time adjustment in the student’s academic pathway” (p.190). Although some researchers may still prefer same-age contrasts, there is considerable justification that same-grade comparisons are probably more appropriate to assess the academic effect of making students repeat a grade.

SCALE OF MEASUREMENT FOR OUTCOME VARIABLES

A methods issue often ignored in the retention literature is the type of scale or metric used to measure student progress. Standard texts in educational and psychological measurement list numerous procedures to quantify scores (for example, raw scores, normed scores, scaled scores, and grade-equivalent scores). Studies of grade retention often use nationally normed tests to compare the progress of retained and nonretained students. However, the substantive implications of using specific measurement scales have been overlooked by most researchers examining the impact of making students repeat a grade. The use of grade-level or grade-equivalent scores to assess academic achievement is particularly problematic. A grade-equivalent norm is basically the average score obtained by students in a specific grade (Thorndike, 1997, p.60). Methodologists have argued that grade-equivalent scores do not allow adequate measurement of change over time between groups. Analyses demonstrate that the measurement scale used to assess academic achievement can lead to divergent conclusions. For example, Seltzer, Frank, and Bryk (1994) compared growth in Iowa Tests of Basic Skills reading scores over the elementary years among students in Chicago Public Schools. Results based on grade-equivalent scores did not measure growth over time as accurately as logit scores derived from a one-parameter item response theory model. Specifically, grade-equivalent scores should not be used to compare students who are in different grades. Coleman and Karweit (1972) contend:

A grade-equivalent score, therefore, means a different thing at every grade level. It does not compare the student to others of the same age or at the same actual grade level; it compares him to the average or median student at *another* grade level....[a grade-equivalent score] is not appropriate for inferences about...*rates of growth* of children at different grade levels (pp. 94-95).

Consequently, it is inappropriate to use grade-equivalent measures when students are the same-age but in a different grade, that is the same-age comparison. Retained and promoted students will be in different grades with the nonretained pupils having an additional year of coursework than the retained. Promoted students will therefore likely score an additional grade-level higher than the retained students who will not have covered the same curriculum. Using grade-equivalent scores to compare the same-grade outcomes of retained and promoted students is less problematic because the amount of curriculum covered will be equal for both the promoted and nonpromoted students.

SAMPLE REPRESENTATIVENESS

To generalise findings regarding the effects of grade retention requires that studies incorporate students from different social and economic backgrounds. If researchers examine only one racial or ethnic group, it is unclear if the effects of holding students back a year will affect individuals in other racial/ethnic categories in a similar manner. Likewise, studies based on students only of European ancestry may yield results inappropriate for students of African American or Hispanic origin who comprise an increasing share of the American educational population. Retention studies should also include economically disadvantaged students as well as those from middle class backgrounds. Some critics of grade retention (for example, Reynolds, 1992) argue that

certain kinds of students (specifically poor minorities) are less likely to be helped by retention than White students of higher social standing.

SAMPLE SIZE AND STATISTICAL POWER

In addition to the problem of the representativeness of student samples, Tanner and Galis (1997, pp.109-110) point out that sample sizes are too small in most retention studies to yield valid conclusions. Sample sizes should be large enough to find potential differences between retained and promoted students. Many samples in retention studies lack sufficient power to reject the null hypothesis of no difference between retained and promoted students. Levels of statistical significance and estimated "effect sizes" can be greatly influenced by sample size. Effect size is often reported to indicate the magnitude of the impact of differences in educational practices. Effect size is often calculated as the difference between the mean outcome for an experimental and control group divided by the common standard deviation of the two groups, or the standard deviation of the control group (for example Holmes, 1989). Oftentimes researchers have only 30 pupils or less in either the retained or control groups. The availability of such a small number of students to assess treatment effects results in very low levels of statistical power. Cohen (1977, p.41) demonstrates that analyses with small sample sizes are not large enough to correctly reject the null hypothesis of no difference even if effect sizes are of a moderate magnitude. For example, assume the true effect size is 0.50 (that is, a difference of one half a standard deviation between two groups), which is often defined as a "medium" effect. The power of a *t* test between the two groups each with 30 observations would be 0.61; the probability of concluding that the two means were the same, when in fact they actually differed, would be 0.39. A sample size of 100 students in each group is required to reduce the probability of making a Type II error to 0.06 if the true effect size is 0.50 (Cohen 1977, p.37). Smaller effect sizes require even larger sample sizes to identify statistically significant differences.

In sum, there are various indicators of the quality of research pertaining to grade retention. The publication outlet, the quality of controls to adjust for preretention differences between retained and promoted students, the basis of comparison between the two groups, and the metric of the outcome measures of student achievement. These characteristics of each retention study will be noted and discussed.

PROMINENT META-ANALYSES OF GRADE RETENTION

The aforementioned criteria are used to evaluate the quality of individual research studies cited in the two most prominent reviews on the effectiveness of grade retention. Holmes (1989, hereafter Holmes) updated his earlier review of retention research (Holmes and Matthews, 1984) and aggregated findings from 63 separate studies. Jimerson (2001) reviewed 20 additional articles examining the effects of grade retention on student outcomes. Whereas Holmes reported largely on research conducted between 1960 and 1987, the studies cited by Jimerson (2001) were published during the 1990s.

As noted above, Alexander et al. (1994) and Dunkin (1996) maintain that published articles are usually of a higher quality when compared to unpublished papers and graduate theses or dissertations. Given that Jimerson (2001) also adopted this suggestion, the current paper focuses only on published research findings from journals and books. Ignored in the 1989 Holmes review are 18 master's theses, 22 dissertations, and two convention papers.² These 42 excluded studies comprise 67 per cent of the 63 studies cited in the Holmes meta-analysis. Only one third of the

² Holmes (1989) cited a convention paper which was later published as a journal article (Shepard and Smith, 1987). Findings from the published paper are presented.

retention studies in the Holmes review were journal articles. Difficulty in acquiring the graduate student documents and convention papers also precluded their inclusion in this review.

The number of studies was further limited by examining only those articles which had findings pertaining to academic achievement. Various researchers have attempted to evaluate the effects of retention on student self-esteem and other psychological outcomes, but the majority of them did not incorporate initial indicators of subjective outcomes (for example, Anfinson, 1941). Given that differences in psychological outcomes between retained and promoted students may have existed prior to the time of retention, it is not possible to assess the impact of making students repeat a grade from many of the studies in the Holmes meta-analyses. Hence, only academic outcomes are considered here: Studies focusing exclusively on subjective dependent variables are ignored. Finally, it was not possible to locate several citations published prior to 1937. These selection criteria left 10 published articles or 16 per cent of those cited in the Holmes (1989) review and 18 out of 20 articles from the Jimerson (2001) meta-analysis of retention studies. The specific articles examined from each of the two meta-analyses are listed in the References.

Holmes Meta-Analysis

Holmes (1989) divided his review into several sections. He first summarised the number of positive and negative differences between the mean outcome measures for retained and promoted students across 63 available studies. He then presented “effect sizes” (mean differences between the retained and promoted pupils divided by the standard deviation of the promoted group) for 25 studies in which the retained and promoted students had been presumably matched or were deemed more similar than in the other 38 studies.³ Only ten of the 63 studies were refereed published articles. Presently no one has rigorously examined the articles in the Holmes meta-analysis. Dunkin (1996, p.95) commented that “Surely, readers of these works [meta-analyses] cannot go to the trouble of the detailed scrutiny required to check the validity of every synthesis before they decide whether or not to rely on them.” The present paper performs such a service by informing readers of the strengths and shortcomings of the articles in the Holmes review. While a discussion of the separate articles may seem tedious, given the frequency that the Holmes meta-analysis is cited (for recent examples see Jimerson, 2004; Shepard, 2004), it is important that readers have some familiarity with the nature of the specific studies in order to determine for themselves the reasonableness of the conclusions inferred from this seminal publication.

Table 1 summarises the major features of the ten published articles cited in the Holmes (1989) meta-analysis. For each article the author and year of publication are listed first. Underneath each reference is the number of students who were required to repeat a grade. The second number specifies the number of promoted pupils in the group whose academic performance is contrasted with the retained students. Several studies include low performing children in addition to students not experiencing academic difficulties and promoted on a regular basis. Occasionally the numbers of students in these additional comparison groups were unspecified and are denoted by a question mark. The other major characteristics for each article are the quality of quality of controls, the presence of a pretest measure of the outcome variable, an initial measure of student ability, the presence of statistical controls for either an outcome measure or level of initial ability, the nature

³ Professor C. T. Holmes (personal communication, April 16, 2004) graciously provided me the name of each study referenced with a number in Table 3 of his chapter on grade retention (Holmes, 1989, p.27). Of the 25 studies Holmes listed as having “matched” subjects, only four were published articles which met the criteria for inclusion in this study. The four studies examined in the current paper which Holmes classified as being of a higher research quality were those by Dobbs and Neville (1967), Peterson, DeGracie, and Ayabe (1987), Sandoval and Fitzgerald (1985), and Shepard and Smith (1987). Nine of the 25 “matched” studies were dissertations, eight were masters’ theses, one was an unpublished convention paper, and three articles examined only psychological outcomes.

of the comparison strategy, the use of grade-equivalent scores, and the authors' conclusion regarding the overall usefulness of retention.

Table 1. Characteristics of Retention Studies in Holmes (1989) Meta-Analysis

Author(s)	Quality of Controls	Outcome Pretest	Ability Pretest	Statistical Control for Outcome Pretest	Statistical Control for Ability Pretest	Type of Comparison	Grade Equivalent Units	Authors' Conclusion Effect of Retention
Kamii and Weikert (1963) (31-33)	Inadequate	No	No	No	No	Same- Grade	No	Negative
Sandoval and Fitzgerald (1985) (32-30-75)	Inadequate	No	No	No	No	Same-Grade	No	Negative ^b
Abidin et al. (1971) (85-43)	Inadequate	No	Yes ^a	No	No	Same-Grade	Yes	Negative
Leinhardt (1980) (44-31-?)	Inadequate	No	Yes ^a	No	Yes	Same-Grade Same-Age	No	Negative ^b
May and Welch (1984) (62-59-102)	Inadequate	No	Yes ^a	No	No	Same-Grade	No	Negative ^b
Shepard and Smith (1987) (40-40)	Inadequate?	No	Yes?	No	No	Same-Grade	Yes	Negative
Niklason (1987) (40-62)	Adequate?	Yes	Yes ^a	Yes	No	Same-Age	Yes	Negative
Dobbs and Neville (1967) (24-24)	Adequate?	Yes?	Yes?	No	No	Same-Age	Yes	Negative
Chansky (1964) (30-33)	Adequate	Yes ^a	Yes ^a	Yes	Yes	Same-Age	Yes	Positive
Peterson et al. (1987) (106-104)	Adequate	Yes	No	No	No	Same-Grade Same-Age	No	Positive

^a Retained and promoted students were significantly different on this variable at time of retention.

^b Data indicate that retention had a positive effect on academic achievement.

Quality of controls is denoted as a dichotomous variable as either “adequate” or “inadequate.” If the article indicated the retained and promoted students were statistically similar on either an outcome measure at the time of retention or an initial indicator of student ability (for example mental ability) the control procedure was classified as “adequate.” Alternately, if the retained and matched promoted students were significantly different on either a measure of initial ability or the initial outcome measure, the quality of the matching is considered to be inadequate. The study is also labelled “inadequate” if indicators of initial ability were available showing retained and promoted students were significantly different and these differences were not statistically controlled. In several articles it was not possible to ascertain if the nonpromoted and promoted students were comparable at the time of retention. Such situations are denoted by a question mark. Of the ten articles, only four meet the criteria for a properly controlled study, and in two of these articles the research designs are questionable. In six of the studies it is highly unlikely that retained and promoted students were similar at the time of retention. Studies can have an adequate research design but still be viewed as “weak” in their ability to assess the effect of grade retention. Small sample sizes, unrepresentative samples, inappropriate grade comparisons, and an inappropriate metric for outcome measures can result in findings with only limited value regarding the impact of retention. An overview of the research characteristics of the ten articles cited in the Holmes meta-analysis are given in Table 1.

One of the first features in Table 1 to raise questions about the extent to which findings from the Holmes meta-analysis can be generalised is the disconcertingly small number of students in the retained and control groups. Only the Peterson, DeGracie, and Ayabe (1987) study examined more than 100 students in the retained and promoted groups. Their analyses, however, were conducted on three separate grade levels (first, second, and third) with samples, respectively, of only 65, 26, and 15 pupils. Small sample sizes of 60 students or less in each group limit the kinds

of conclusions which can be made. Note that Chansky (1964) as well as Dobbs and Neville (1967) made inferences based on sample sizes of 30 pupils or less. Admittedly many educational researchers have access to schools with a limited number of students. While investigators can only use data which are available; it is unwise to make broad generalisations about the effectiveness of retention practices based on small numbers of observations. Further, such small samples make it highly unlikely that researchers will be able to find statistically significant differences between retained and socially promoted students. Sampling variation alone could account for the findings. Since the mid-1960s economists and sociologists have required that nonexperimental research examining individual student achievement be based on representative samples with at least several hundred students, if not more, before the findings are taken seriously. Not one of the cited articles meets current standards of sample size requirements for acceptable nonexperimental social science research.

As is evident, another aspect of these articles limiting their generalisability is that the samples are based on students who attended only a small number of schools in but a few selected school districts. While students from different geographic regions are represented across the different articles, pupils in each study are from only one state. None of the studies provide a student sample representative of the state school population. Were one to require that an adequate research design include a sample size with sufficient statistical power and students who were representative of all pupils enrolled in schools, none of the ten published articles from the Holmes meta-analysis would meet the criteria of an acceptable study. Insofar as some readers may find such criteria too stringent, other characteristics of these studies are examined.

Although Table 1 shows that the authors in eight of the ten studies conclude that making students repeat a grade does not have a positive effect on academic achievement, the overall quality of the research is highly suspect due to inadequate research designs. Table 1 indicates that in only two studies were the retained and promoted students likely comparable in ability or educational progress at the time of retention (Dobbs and Neville, 1967; Peterson et al., 1987). A major problem in interpreting findings from the remaining eight articles is that the nonpromoted and promoted children were of different ability levels at the time of retention, or it was not possible to determine their initial academic progress. Authors were often unable to control for possible differences between retained and promoted students that existed prior to retention. Two studies (Kamii and Weikart, 1963; Sandoval and Fitzgerald, 1985) provided only student outcome measures; no prior information was available to indicate if the retained and promoted students differed in academic ability when students were required to repeat a grade. Kamii and Weikart (1963) examined the educational performance of Michigan seventh graders. Data were obtained retrospectively from the school records of 31 students who had been retained once between first and fifth grade and 31 pupils in the same school district who had never been required to repeat a grade. Measures of student achievement were available only in seventh grade. Analyses using *t* tests revealed that in seventh grade the retained students obtained significantly lower mean reading and arithmetic scores on the Iowa Test of Basic Skills, significantly lower course grades, and significantly lower IQ scores when compared with the promoted students. Although the standard interpretation of such findings is that grade retention does not help students, such a claim is questionable because the promoted students may have demonstrated higher levels of achievement at the time the other students were retained. The lower mean seventh-grade IQ scores of the retained students suggest that they may have had lower levels of mental ability than the nonretained pupils before retention.

Holmes classified the Sandoval and Fitzgerald (1985) article as being one of the “better matched” studies, but their research design is clearly inadequate. The authors stated that they had matched 75 grade repeaters with students of “comparable ability and motivation” since both sets of students were in the same high school classrooms (Sandoval and Fitzgerald, 1985, p.166). To infer that these California high school students were cognitively similar is highly questionable as

students in the same classroom can vary considerably in ability. Retained students in this study had been held back beginning in kindergarten through any other grade prior to high school. To correctly assess the impact of grade retention requires that retained and promoted pupils have similar levels of ability at the time of retention, not years later.

Four other articles in Table 1 designated as having inadequate research designs also provided no initial measures of the outcome indicators between the nonpromoted and promoted students (Abidin, Golladay, and Howerton, 1971; Leinhardt, 1980; May and Welch, 1984; Shepard and Smith, 1987). However, these studies attempted to match student cognitive levels by using tests of students' readiness to learn. Abidin et al. (1971) attempted to control for the ability of retained and promoted students in a south-eastern American school district by matching on the Metropolitan Readiness Test administered to pupils in first grade. Leinhardt (1980) used results from the First Grade Screening Test to identify students for a control group to be compared with retained students. Developmental scores from the Gesell Screening Test were the basis of May and Welch's (1984) identification of retained and promoted kindergartners with similar levels of ability at the time of retention. Shepard and Smith (1987) attempted to match retained and promoted Colorado kindergartners with similar initial levels of developmental ability using the Santa Clara Readiness Inventory. In each of these four studies, however, there is evidence that the nonpromoted and promoted students in the control groups were not of equal ability. Holmes classified the Shepard and Smith article as being one of the better matched studies, but the authors formed matched pairs of children from readiness scores obtained at the *beginning* of the kindergarten year in September, not the *end* of the year when retention decisions were made. The promoted students could have progressed more rapidly than the retained children during kindergarten, resulting in groups of unequal levels of school readiness at the end of the school year. Shepard and Smith (1987, p.350) further acknowledged that 11 of the 40 matched pairs of students exhibited dissimilar scores on the Santa Clara Readiness test.⁴ May and Welch mentioned that the 62 retained kindergartners in a suburban New York state district were more developmentally immature than the control group of 59 children, also recommended for retention but promoted. Neither Shepard and Smith nor May and Welch used the readiness scores as covariates to control for the differences in ability levels between the retained and nonretained children. For these reasons, the research designs of these two studies are classified as inadequate. A more detailed discussion of these four articles follows.

Findings from the Abidin et al. (1971) study are extremely problematic to interpret because of the lack of similarity between promoted and nonpromoted pupils. Although Holmes did not classify the Abidin et al. article as one of the better controlled studies, this paper illustrates the difficulty in reaching a conclusion about the impact of retention when retained and nonretained students with different characteristics can not be properly matched. The authors attempted to equalise the initial ability of retained and promoted children in a single school district by matching on scores from the Metropolitan Readiness Test (MRT) taken by first grade students. The academic ability and performance of 85 pupils who had been retained either in first or second grade were compared with 43 children scoring in the bottom quartile of the Metropolitan Readiness Test who had never been retained. Findings from *t* test comparisons between the low-performing promoted and retained students showed the former obtained higher scores on the fourth- and sixth-grade reading and mathematics sections of the SRA. Given that the retained students evidenced higher levels of readiness for school than the nonretained students, the authors concluded that holding students back a year was a "noxious" educational policy. Because many critics of retention (for example, Jimerson, 1999; Meisels and Liaw, 1993; Niklason, 1984; Reynolds, 1992; Southard

⁴ Shepard and Smith (1987) stated that unreported analyses based on the 14 pairs of students who were exactly matched on the Santa Clara Readiness Inventory resulted in effect sizes similar to those presented in the paper. However, it is important to note that the retained and promoted students may not have been similar in ability because they were matched prior to the time of retention.

and May, 1996) cite Abidin et al. as demonstrating the failure of making students repeat a grade, it is worthwhile to thoroughly examine this study.

The data do not support the strong conclusion Abidin et al. (1971) assert. First, retained students and promoted students differed on other characteristics which were uncontrolled. Boys, African Americans, pupils of lower economic standing, and students with working mothers were more likely to be retained, but these variables were not controlled. Other information further indicated the promoted and nonpromoted students were not comparable. Although the authors reported that teachers estimated the retained and nonretained students should have similar academic success, the great dissimilarity in Lorge-Thorndike IQ scores between the two groups suggested that the retained and promoted students were not adequately matched on intellectual ability in first grade. The authors mentioned that the promoted students had significantly higher IQ scores on both the fourth- and sixth-grade tests. Unfortunately no IQ scores were available at the end of first grade, the time retention decisions were initially made for the majority of students, to determine if the retained and nonretained children were of comparable intellectual ability. If one assumes that intelligence levels remain fairly stable over time, the IQ scores of the retained children were likely lower in first grade than the scores of the promoted pupils. The greater intellectual ability of the nonretained children is probably the cause of their higher academic scores in the fourth and sixth grade rather than not being required to repeat a year in grade as Abidin et al. (1971, p.415) claim.

Another factor which limits the interpretability of the Abidin et al. (1971) findings is the reason for grade retention. Besides academic ability, first grade students were retained because of behavioural problems and infrequent school attendance. The observed larger first-grade readiness scores of the retained students may have occurred due to the fact that 15 per cent of the retained pupils were held back because they had not accumulated enough school days to be promoted. It is not unreasonable to speculate that some students who were held back because of excessive absences (perhaps due to illness) could have had higher initial readiness scores than the lower-scoring promoted pupils who did not miss as much school. Deleting from the analyses those students retained for reasons other than academic ability would have helped clarify the relationship between retention and later school success.

The fact that 40 per cent of the retained students (34 of the 85 retainees) were held back in second grade further compromises the matching process because the retained second graders likely evidenced higher first-grade readiness scores than the first graders who were required to repeat that grade. Thirty two per cent of the students repeating the first or second grade were held back because of academic failure, but readers are not informed if the grades of the retained first graders were lower than those of the promoted first graders. Instead, the authors stated that the grades of the retained and promoted groups were the same in first grade.⁵ Similarity in subject matter grades in first grade occurred probably because the grades of those students retained later in second grade were comparable to those of the promoted first graders. Combining students retained in first or second grade makes it difficult to gauge the impact of grade retention. A more accurate assessment of retention could have been made if the retained group consisted only of first graders who were held back for academic but not emotional or behavioural reasons.

The four previously mentioned articles tried to match students on the basis of school readiness tests, but research has questioned the usefulness of such tests. Although it seems reasonable to assume a positive correlation exists between school readiness tests and later student academic

⁵ The degree to which the grades of retained and promoted students in first grade differ is confusing. Abidin et al. (1971, p.414) also state: "However, the retained group's grades in both reading and mathematics were significantly lower the year they were retained in the first grade. During the second and third grade there were no significant differences in grades between the promoted and retained group nor for the retained group between grades repeated and the original grades."

ability, it is unclear whether readiness tests adequately predict how well students will perform in school. Readiness tests are often so basic (for example, does the student know to read from right to left; can the child recognise certain sounds; does the pupil know major letters, etc.) that they may not be strongly related to academic ability or general intelligence. Several measurement psychologists have argued that the Metropolitan Readiness Tests (MRT) used by Abidin et al. (1971) is not a valid measure of reading ability. Stoner (1995) cited research indicating that the skills assessed by the MRT have relatively little relationship with the ability of beginning learners to read. Mabry (1995) is even more critical of the MRT and stated: "The test is outdated, passé in terms of learning theory, technically inadequate, confusing to targeted audiences, and likely detrimental to children and schools" (p.612). These sentiments may explain why recent studies seldom report use of the Metropolitan Readiness Tests to gauge student ability. The adequacy of the Gesell Test of school readiness used by May and Welch (1984) has also been challenged. Although extensively used in earlier decades, the Gesell tests have significant shortcomings; for example, g., uncertain reliability and validity, inadequate norming samples, and an inability to predict future performance (Kaplan and Saccuzzo, 1997, pp.307-309). Information describing the measurement properties of the Santa Clara Readiness Inventory used by Shepard and Smith (1987) to match retained and nonretained students could not be located.

Unlike the other three studies (Abidin et al., 1971; May and Welch, 1984; Shepard and Smith, 1987) which attempted to control only by matching students on level of school readiness, Leinhardt (1980) used the First Grade Screening Test readiness score in an analysis of covariance to statistically adjust for initial differences in ability levels between retained and promoted children at the end of their kindergarten year. There were 44 pupils whose scores resulted in their being placed into a transition first-grade room (another form of retention) rather than being promoted directly to first grade. The control group consisted of 32 pupils from the previous year's kindergarten class whose screening scores were below the level recommended for promotion to first grade: Nonetheless, these children had been promoted to the first grade. Students in the transition room received individualised reading instruction, as did nine of the 34 at-risk pupils who had been promoted to first grade instead of being placed in the transition room. However, the majority of promoted children experienced a different type of reading program; 23 of the 32 promoted students received basal instruction while all 44 pupils placed in the transition room received New Reading System (NRS) instruction. The screening score and type of reading program were used as covariates to assess the effect of retention on reading performance (measured by the Stanford Achievement Test). Regression results indicated that the retained pupils performed 19 points lower than the promoted kindergartners at the end of 1977 when the promoted students had completed first grade and the nonpromoted pupils were at the end of the transition year. However, by the end of their first grade year, the reading scores of the transitioned students were comparable with those of their first grade classmates who had been promoted to first grade after kindergarten. The first-grade SAT reading score for the retained students was 77.9 while the end of first-grade reading score of their younger classmates was 60.5, a difference in reading scores of 17.4. This gap was statistically insignificant, probably due to the small sample, but the effect size of 0.97 was large.⁶ Had Leinhardt used the initial prescreening test score as a covariate, the difference in first grade reading results between retainees and their year younger classmates might have been even larger. Nonetheless, Holmes (1989, p.26) as well as Heubert and Hauser (1999, p.120), contend that Leinhardt's study demonstrates the ineffectiveness of grade retention.

⁶ Holmes (1989, p.18) computed the effect size by dividing the mean difference between the retained and promoted students by the standard deviation of the promoted group. For this specific situation, the effect size was calculated as follows: $(77.9-60.5)/17.9=0.97$. Unless specified otherwise, the same computational procedure is used to estimate effect sizes throughout the paper.

However, opponents of grade retention ignore Leinhardt's (1980, pp.59-60) discussion which reveals that the major discrepancy between the transition-room pupil reading scores and the first-grade reading scores of the promoted pupils most likely resulted from differences in the quality of reading instruction between the two groups. Pupils in the 1976-1977 transition room were given two and one-half hours *less* reading instruction per week than the promoted kindergartners who were in first grade that school year. In addition, the transitioned students received *less* than one-half of the test-relevant information given the promoted pupils. The promoted students were also tested more frequently than the kindergartners put in the transition room. The promoted students "were taught the basics of reading directly, more often, and for longer periods of time" (Leinhardt, 1980, p.60). Instead of being given direct instruction in reading, pupils placed in the transition first grade were taught learning skills. The extra help retained students received was inferior to that provided the promoted pupils. Insofar as placing students in a transition room was associated with poorer quality reading instruction, it is not possible to reach any valid conclusions about the effects of a transition first-grade year from this study. The Leinhardt study is inconclusive because the treatment condition (retention) is completely confounded with the nature of reading instruction.

The quality of the research design from the Niklason (1984) study is labelled adequate, although the degree to which initial differences between retained and promoted children were controlled is uncertain (as indicated by a question mark in Table 1). Unlike most of the studies cited by Holmes, Niklason included both an initial indicator of the outcome measure as well as student intellectual functioning in her analysis of urban and suburban Utah elementary school students. Teachers recommended that 144 pupils repeat their current grade (kindergarten through sixth grade). Among the 102 students with available data during the period of study, only 40 students were retained in spite of the teachers' recommendations. Both groups of students were tested at the end of the school year in 1980 and 1981. The 1980 score for either reading or arithmetic from the Wide Range Achievement Test was used as the covariate to predict the 1981 score. Niklason reported promoted pupils showed significantly greater growth in reading achievement than the retained students between 1980 and 1981. The author concluded: "The results of this study...showed that retaining students did not serve the intended purpose of increasing the student's growth academically or in personal or social adjustment" (Niklason, 1984, p.496). This interpretation is problematic because the author also reported that in 1980 the retained students had significantly lower levels of performance ability, along with greater problems in personal and social adjustment, than the promoted children. Unfortunately, Niklason never defined the term "performance ability" but it does not appear to be the same as verbal intellectual ability or academic achievement. Had Niklason included earlier measures of performance ability, personal, and social adjustment as additional covariates, along with the pretest measure to control for these student differences, the findings would have been more conclusive.⁷

Similar to Niklason (1984), two other studies included indicators of student ability and an initial outcome measure, but these two studies may have better controlled for initial differences between retained and nonretained students. Dobbs and Neville (1967) appeared to better match students on initial levels of academic achievement and ability. They compared 30 first graders who had to repeat the grade with 30 never-retained second graders. Pupils were matched on race, sex, type of classroom, age, mental ability, reading achievement, and school socioeconomic status. Arithmetic

⁷ In a later paper Niklason (1987) attempted to control for variables other than the pretest value of the outcome measure. Instead of using the control variables as covariates, however, she created dummy variables and performed an analysis of variance which reduced the power of the control variables. This reanalysis of the Utah data is also problematic because the author used a hierarchical procedure to gauge the effect of grade retention which appears not to have controlled for the other variables later entered into the equation (Niklason, 1987, p.342).

scores of the retained students were significantly behind those of the promoted. Beginning in 1962 with reading and arithmetic scores from the Metropolitan Achievement Test, additional test results were obtained at the end of the 1963 and 1964 school years among the remaining 24 students with complete data. The promoted pupils showed significant greater gains in both reading and arithmetic than the retained students. Given the lack of statistically significant differences between the promoted and nonpromoted with respect to age, IQ, and reading level, Dobbs and Neville concluded “that promotion led to the increased achievement gain of the promoted group” (p.474).

Still, the comparability of retained and never-retained students is unclear because the authors did not state if the reading scores for the retained students were taken at the end of their initial year or repeated year of the first grade. Likewise, the test grade year used to match the promoted students with the retained is vague. One presumes that the 1962 achievement scores used to match the two groups of students were the initial first grade test results for both the retained and promoted pupils. Although the initial group means for mental ability and reading achievement were statistically similar, these comparisons were based on 30 students in each group. However, the analyses assessing the change in grades over time were based on only 24 students, as some pupils initially in the study had missing data in later years. A difference of six students may seem minor, but 20 per cent of the initial sample was missing. No information was given indicating if the matched samples of 24 students actually analysed over time were similar in IQ or reading ability. The fact that the IQ difference between the initial 30 retained and promoted students approached statistical significance ($t = 1.50$) suggests that the promoted pupils had higher mental abilities than the children required to repeat first grade. In addition, students may not have been well matched on socioeconomic status because matching was based on school-level rather than individual-level characteristics. Not all children within a school necessarily have families with similar income and social characteristics. These features of the study raise questions about the overall quality of the research. The degree of similarity between the retained and promoted students is uncertain.

Of the two studies reporting that holding students back a grade had a positive effect on academic performance, Chansky (1964) utilised a different statistical adjustment procedure to equalise differences in ability between nonpromoted and promoted low performing children. Thirty first-graders, whose teachers considered the students to be a “poor risk” for success in second grade, were held back, while 33 low-performing students who were judged by their teachers and principals to be a “good risk” were promoted to second grade. First-grade teachers in a rural New York county initially selected 63 students to be retained but 30 were placed into second grade. The promoted pupils scored significantly higher on the California Achievement Test and also evidenced higher levels of mental ability than those retained at the end of the study period. However, the California Achievement Test enabled computation of an Intellectual Status Index which measured academic performance relative to mental age. Although the promoted students showed slightly greater gains in reading by the end of the next year, comparing observed performance *relative to expected performance*, indicated that “...the retained group was less discrepant from mental age expectancy than the promoted group” (Chansky, 1964, p.230). Requiring the more academically challenged students to repeat first grade at least enabled them to perform at a level consistent with their mental ability. The promoted students, however, had much lower scores in reading comprehension, arithmetic reasoning, and arithmetic fundamentals than expected, given their intellectual ability. The observed positive effect of retention may partially be a function of the measurement procedure, as this is the only study using a discrepancy technique to control for differences in the mental ability of retained and promoted students.

The only other published study, appearing in the Holmes review, reporting a positive effect for retention examined elementary school students in Mesa, Arizona. Holmes listed this article as having one of the “better matched” groups of students. Given that the retained and nonretained

students exhibited similar scores on the outcome measures at the time of retention, the study also met the criteria of adequate controls. Peterson et al. (1987) tracked the academic performance of first, second, and third grade students who repeated a grade. Students were matched on sex, ethnicity, age, and scores from the California Achievement Test (CAT). Separate analyses, by grade, over a three year period revealed that the retained students obtained higher reading, language, and math scores than their promoted counterparts. Although the advantage of first grade retention no longer existed in reading by third grade and mathematics by second grade, the benefits to second and third grade retention persisted over the next three years. The authors concluded that making students repeat a grade resulted in positive academic achievement.⁸

Thus far only four studies have been identified from the Holmes articles which can be classified as having adequate controls for differences between retained and promoted children (that is, Chansky, 1964; Dobbs and Neville, 1967; Niklason, 1984; Peterson et al., 1987), although the assessment is questionable in two of the articles. Holmes did not list either the Chansky (1964) or the Niklason (1984) analyses in his table of "better matched" studies. Whereas the findings in Chansky (1964) and Peterson et al. (1987) favoured making students repeat a grade, the other two studies concluded that retention was unsuccessful in improving student performance. It appears as though the four studies with better controls yield conclusions which, at best, cancel out one another. However, additional questions can be raised about the methodological quality of the two negative studies. Differences between the retained and socially promoted students may not be as well controlled in the Dobbs and Neville (1967) and Niklason (1984) articles as occurs in the other two articles. Further, the two articles with negative conclusions were biased against finding a positive effect for retention given the type of comparisons made. Dobbs and Neville (1964), as well as Niklason (1984), compared retainees and nonretainees when they were in different grades (that is, the same-age contrast). These authors ignored that the scores of the retained students were one grade behind the promoted children, who had experienced an additional year of new curriculum than those held back. The higher scores of the socially promoted students may derive from their having an additional year of being introduced to more advanced material than the retainees. Another feature of the two analyses favouring the promoted pupils was that the outcome measures were based on grade-equivalent units. Given that they had not been introduced to the next year's curriculum, the retained students were less likely to score as high as the promoted students, who were one grade ahead and had received an additional year's worth of instruction of more advanced educational material.

Although only two of the ten reviewed articles from the Holmes (1989) meta-analysis concluded that making students repeat a grade was worthwhile, findings from several of the other articles can be interpreted to support retention. For example, even though they believed retention in later grades was ineffective, Sandoval and Fitzgerald (1985, p.169) stated that children who attended a first-grade transition room between kindergarten and first grade, rather than being promoted, demonstrated academic performance superior to their "matched" senior high classmates. Proponents of retention argue that making students repeat a grade will help them learn the material needed to catch-up with classmates who did not have academic difficulties. Two articles

⁸ Alexander et al. (2003, p. 24) suggest that the higher test scores of the retained students relative to the nonretained children may partially be due to the use of a measurement scale based on normal curve equivalents (NCEs). Whereas the means of the retained first graders were based on NCEs normed to first grade students, the means of the second graders would have been scaled relative to the distribution of all second graders used to calculate the national norms. In short, the means of the retainees and the nonretained students were not based on the same reference group used to norm their raw scores. This lack of a common reference point may be problematic for Peterson's et al. (1987) same-year analyses but the same-grade mean contrasts are likely based on norms derived from the same set of pupils. For example, the first grade means of both the retained and promoted pupils are probably derived from the same set of first graders throughout the nation used to establish the first grade norms.

indicated that students who had been struggling with subject matter learned enough after a year of retention to reach a level of academic achievement comparable with nonretained peers. Leinhardt (1980) reported that the retained students had higher first-grade Stanford reading scores than their classmates who had not been placed in a transition room. May and Welch (1984) found that kindergartners who had been required to repeat a grade because of low-developmental readiness scores later performed as well as other developmentally immature students who had been socially promoted. The mean third grade New York Pupil Evaluation Test scores were similar between the retained and socially promoted, as were their scores on the Stanford Achievement Tests. This study concluded retention was ineffective because the authors assumed the retained students should have had higher scores than their socially promoted peers who had experienced one less year of school. But May and Welch overlooked the fact that the retained students were initially significantly less mature than the socially promoted students (p.384). One could assume that a transition year helped the lower-performing children catch up with their nonretained peers. The absence of a statistically significant difference between promoted and nonpromoted pupils can be interpreted to support an advantage for retention if the performance of the latter students was appreciably below that of the nonretained children at the time of retention.

Critics of grade retention often cite mean “effect size” differences between retained and nonretained students to conclude that making students repeat a grade is ineffective. Holmes (1989, p.27) reports an average effect size of -0.28, based on the 25 studies he classified as “best-matched. Even if only the 16 studies which had IQ or achievement scores which presumably measured initial student ability are considered, the effect size is -0.30. However, the term “effect size” is a misnomer, particularly in the retention literature (and probably in other substantive areas), because *effect* denotes a causal change or the impact of a treatment on an outcome measure. It must be remembered that an “effect size” is merely a mathematical term indicating only how many standard deviation units two means differ. The important question is not the average magnitude of the differences, but what are the factors which *explain* the observed differences? Effect size can not be interpreted to imply retention causes an outcome, unless alternative explanations can reasonably be ruled out. The average effect size which can be calculated from the higher quality studies in Holmes varies by which studies are included in the analyses. For example, Shepard, Smith, and Marion (1996, p.252) reported the average effect size from the “six most tightly controlled studies showed retained students behind the controls by one-quarter of a standard deviation.”⁹ However, if the Peterson et al. (1987) study, which controls for initial achievement level, SES, and other characteristics, is included with the six studies used by Shepard et al. (1996), the average effect size is reduced in half to -0.10.

These calculations indicate that the choice of variables to be controlled also influences the estimated average effect size. Assume that one were to predict an academic outcome by controlling for initial scores on an achievement test along with socioeconomic status. Holmes (1989, p.27) listed only five studies among the 25 studies with students considered matched on both initial achievement score and SES.¹⁰ The average effect size for these studies is a greatly reduced -0.01. Thus there is no meaningful difference between retained and promoted students after taking into consideration initial level of academic achievement and social-economic

⁹ Although Shepard et al. (1996) did not specify which of the studies listed in Holmes were the “six most tightly controlled,” the average effect size from the Anfinson (1941), Archer (1967), Dobbs and Neville (1967), Schuyler and Matter (1983), Shepard and Smith (1987), and Wright (1979) studies is -0.24. The degree to which some of these studies control for initial differences is questionable (e.g., Dobbs and Neville, 1967; Shepard and Smith, 1987). In particular, Anfinson had no measure of student social and personal adjustment prior to retention.

¹⁰ The five studies which attempted to match on achievement test and SES are listed here in order of study number Holmes provided. Following the study number is the reference to the study and the average computed effect size listed in Holmes (1989, p.27): #7=Dobbs and Neville (1967) -0.63; #17=Peterson et al. (1987) 0.76; #19=Schuyler and Matter (1983) -0.41; #20=Shepard and Smith (1987) -0.13; #25=Wright (1975) 0.35.

background. If the Dobbs and Neville (1967) article were deleted from the computations because of the uncertainty of the matching process, the average effect size would be 0.12, indicating that retention results in weak positive academic benefits. In sum, the impact of grade retention on student performance depends on which variables are considered more important as controls and how well retained and promoted students are actually matched.

The indiscriminate use of these reported effect sizes would lead one to argue that retention does no good, but such an interpretation is valid only if the two sets of students were similar at the time of retention. As has been pointed out, it is unclear whether retained and socially promoted students in the Dobbs and Neville (1967) or Shepard and Smith (1987) study were truly similar in ability at the time of retention. There are other variables that should be controlled in the cited analyses. For example, retained and nonretained students may differ with respect to parental support. Several studies (May and Welch, 1984; Niklason, 1984; Shepard and Smith, 1987) noted the low-performing students in the control groups were often socially promoted at the request of their parents. Such parents may be more concerned about their children's education and provide them additional academic assistance that retained students do not receive from their parents, but these potential family differences are usually ignored. The matching procedures are so inadequate in most of the citations provided by Holmes that few studies yield valid conclusions, regardless of how individual authors summarise their findings.

Although critics of grade retention extensively cite the Holmes (1989) review as definitive evidence against making students repeat a grade, they overlook Wilson's (1990, p.229) critique chiding Holmes for making "glaring mismeasurements and misinterpretations...." The major problem with most of the articles Holmes cited is their inability to adequately control for initial levels of academic achievement and ability between retained and promoted students. Even in most of the studies Holmes classified as better matched, retained and promoted students were dissimilar on initial levels of academic ability: Very few of these articles can be considered to be "tightly controlled." Consequently, conclusions from the Holmes meta-analysis about the impact of making students repeat a grade are not persuasive as many claim. One could argue that the research designs were better in the unexamined unpublished dissertations and masters' theses, but this is unlikely. Had the matching procedures or research methods been comparable to that used in the published articles, one would assume the graduate papers would have been published or appeared in some other outlet. It is more probable that the problems of inadequate matching, lack of statistical controls, improper comparisons, and small sample size also pervade the dissertations and masters' theses cited by Holmes. Contrary to the assertions of many authorities, an unbiased reader should conclude that there is no overwhelming body of evidence in the Holmes (1989) meta-analysis to support the contention that grade retention is an ineffective or harmful educational remediation strategy. Many of the shortcomings found in the Holmes review are also prevalent in Jimerson's (2001) synthesis of more recent studies.

Jimerson Meta-Analysis

Jimerson (2001, hereafter Jimerson) published the most recent comprehensive review of grade retention studies. His meta-analysis contained 20 refereed articles, but only 18 examined academic achievement. Unlike Holmes, Jimerson made no effort to differentiate which of the studies in his synthesis utilised better-matching or statistical controls to account for initial differences between retained and promoted students. The indicators of research quality previously applied to the Holmes meta-analysis were also used to scrutinise Jimerson's review. First examined are the 16 studies in which the authors concluded that grade retention had no positive effect on student academic achievement. Similar to the review of the Holmes articles, studies cited by Jimerson are described in order of their rigor as determined by adequacy of the research design to control for potential student differences. Less rigorous studies are described first. However, an exception is sometimes made in the order of presentation. Articles based on the

same data are jointly described regardless of the quality of research design. The two studies in which the authors favoured the practice of making students repeat a grade are examined last. Once again a detailed discussion of many articles is given. Although such a listing may seem monotonous, readers need to be aware of the unique features of these articles in order to form their own conclusions about the extent to which making students repeat a grade affects their academic performance. A summary of the major characteristics of the 18 studies from the Jimerson meta-analysis is presented in Table 2.

Table 2. Characteristics of Retention Studies in Jimerson (1989) Meta-Analysis

Author(s)	Quality of Controls	Outcome Pretest	Ability Pretest	Statistical Control for Outcome Pretest	Statistical Control for Ability Pretest	Type of Comparison	Grade Equivalent Units	Authors' Conclusion Effect of Retention
Johnson et al. (1990) (20-17-20)	Inadequate	No	No	No	No	Same-Grade	No	Negative ^b
Hagborg et al. (1991) (38-38)	Inadequate	No	No	No	No	Same-Grade	No	Negative
Meisels and Law. (1993) (3,203-13,420)	Inadequate	No	No	No	No	Same-Grade	Yes	Negative
Dennenbaum and Kulberg (1994) (25-28-17-25)	Inadequate	No	Yes [?]	No	No	Same-Grade	No	Negative
Ferguson (1991) (46-20)	Adequate	No	Yes	No	No	Same-Grade	No	Negative ^b
Ferguson and Mueller-Streib (1996) (33-14)	Inadequate	No	Yes [?]	No	No	Same-Grade	No	Negative
Phelps et al. (1992) (24-24)	Inadequate	Yes ^a	No	No	No	Same-Age	No	Negative ^b
McCombs-Thomas et al. (1992) (31-31)	Inadequate	Yes	No	No	No	Same-Age	No	Negative
Mantzicopoulos and Morrison (1992) (53-53)	Adequate [?]	Yes [?]	Yes ^a	No	No	Same-Grade Same-Age	No	Negative
Mantzicopoulos (1997) (25-15)	Adequate [?]	Yes [?]	Yes ^a	No	Yes	Same-Grade Same-Age	No	Negative
Rust and Wallace (1993) (60-60)	Adequate	Yes	No	No	No	Same-Age	No	Negative ^b
Jimerson et al. (1997) (32-50)	Adequate	Yes	Yes	No	No	Same-Grade	No	Negative ^b
Jimerson (1999) (20-23)	Inadequate	Yes	Yes ^b	No	No	Same-Grade	No	Negative
Reynolds (1992) (231-200-?)	Adequate	Yes	Yes	Yes	Yes	Same-Grade Same-Age	Yes	Negative
Reynolds and Bezruczko (1993) (251-1004?)	Adequate	Yes ^a	Yes ^a	Yes	Yes	Same-Age	No	Negative
McCoy and Reynolds (1999) (315-843)	Adequate	Yes ^a	Yes ^a	Yes	No	Same-Age	No	Negative
Pierson and Connell (1992) (74-69-35-60)	Adequate	Yes	Yes	No	No	Same-Age	No	Positive
Alexander et al. (1994) (242?-106?)	Adequate	Yes	No	Yes	No	Same-Grade Same-Age	No	Positive

^a Retained and promoted students were significantly different on this variable at time of retention.

^b Data indicate that retention had a positive effect on academic achievement.

Comparable to the studies cited in the Holmes (1989) meta-analysis, most of the findings in the Jimerson survey are based on samples limited to only a small number of students. Of the 18 studies, 12 reported fewer than 60 pupils in either the retained or the nonretained control groups. Reynolds (1992) as well as Alexander et al. (1994) tracked the school performance of more than

200 academically challenged children. Meisels and Liaw (1993) were able to analyse the academic achievement of thousands of middle school students from a national representative sample of the U.S school population. With the exception of the Meisels and Liaw study, all of the other analyses were based on students from usually only one school district. The small sample sizes and restricted geographic locations of the studies limit the generalisability of the findings.

A major problem in many of these studies is again the lack of baseline data which would indicate if the promoted pupils were comparable to the retained student when they were held back in grade, an assumption underlying many of the articles Jimerson cited. None of the first four studies listed in Table 2 had indicators of student ability or outcome measures at the time of retention. Johnson, Merrell, and Stover (1990) compared 20 students retained in kindergarten or first grade from four public schools in the state of Washington with 17 pupils recommended for retention, but who were advanced to the next grade. Given that this was a retrospective study, there were no indicators of student ability or academic achievement prior to fourth grade. The outcome measures were fourth grade Metropolitan Achievement Test (MAT) scores. The absence of any statistically significant differences between the retained and promoted pupils recommended for repeating a grade, led Johnson et al. (1990) to conclude "... the use of early grade retention was not effective as an academic intervention..." (p.337), but the authors did not acknowledge that these academic differences may have existed prior to retention.

Hagborg, Masella, Palladino, and Shepardson (1991) concurred that retention was ineffective after examining the academic performance of 38 high school students from a semi-rural New York school district who had been retained prior to eighth grade. The students' Comprehensive Test of Basic Skill scores (taken at the end of eighth grade), along with their Lorge-Thorndike Verbal and Non-verbal IQ scores (obtained at the end of seventh grade), were compared with the scores of students of the same sex who were in the same-track English class. On average, the 38 never-retained students had statistically significant higher scores. Such findings are not unexpected because the retained students were probably very different in initial levels of ability than exhibited by the promoted students who had never been retained. Without knowing how similar pupils were at the time of retention, any inference from this study is questionable.

The retention study cited in the Jimerson meta-analysis with the largest number of students is based on the 1988 National Education Longitudinal Study (NELS). Meisels and Liaw (1993) examined the grade retention histories (as reported by parents) of over 16,000 eighth grade students. More than 3,000 pupils had been retained once between kindergarten and grade eight. Results from reading and mathematics tests, prepared by the Educational Testing Service when the students were in eighth grade, along with self-reported grades were contrasted between the nonpromoted and promoted groups. Retained students were found to have significantly lower grades and standardised test results after controlling for gender, race, and family background characteristics. Meisels and Liaw (1993, p.76) concluded: "This study confirms that retention does not succeed in reducing the risks of later school failure." Such an inference was not substantiated, however, because there were no variables prior to eighth grade which measured earlier student academic ability.¹¹

Three of the studies cited in Jimerson's summary were unable to match students on initial outcome measures, but instead attempted to match students on other indicators of student ability (Dennebaum and Kulberg, 1994; Ferguson, 1991; Ferguson and Mueller-Streib, 1996). Dennebaum and Kulberg (1994) tracked the academic progress of 95 elementary students from a Rhode Island school district. After examining school records of fourth and fifth graders, the

¹¹ Critics of grade retention sometimes incorrectly summarize the nature of earlier studies. For example, McCoy and Reynolds (1999, p.276) report that Meisels and Liaw (1993) "adjusted for prior achievement" between the retained and promoted students in their article. This assertion is false because Meisels and Liaw (1993) clearly state that they were unable to control for academic achievement prior to or at the time of retention.

authors created four groups of students. One group (n=25) had been retained in kindergarten before promotion to first grade. Twenty eight kindergartners were placed into a transition program before being promoted to first grade. A small number of kindergartners (n=17) had been recommended to repeat kindergarten or for placement in a transition room; instead they were advanced to first grade. A control group of 25 kindergartners were promoted directly to first grade. Reading, mathematics, language, and total battery scores from the Metropolitan Achievement Test, Sixth Edition (MAT-6) were the outcome measures for academic performance in first, second, and third grade. To determine if students were of comparable academic ability, mean scores from the Otis-Lennon Ability Test, Fifth Edition were examined at the end of first grade. The authors imply that the students in each of the four groups were comparable in ability because an analysis of variance revealed no statistically significant difference in the School Ability Test scores from the Otis-Lennon Ability Test. Students who had repeated kindergarten were observed to have lower fourth grade scores than the promoted students. Further, pupils placed in the pre-first transition grade also had lower fourth grade scores than the students in the two promoted groups. The authors asserted “the results indicate that retention actually hurt their achievement when compared to the children who were recommended for retention but went onto first grade anyway” (Dennebaum and Kulberg, 1994, p.11). They reasoned that retained students should perform significantly better than their unretained classmates who were a year younger. Although they do not present any data, the authors reported that no significant differences were found between the retained pupils and students in their classrooms who had not been required to repeat a year. The extra year spent repeating kindergarten or enrolling in the transition room was therefore viewed as a waste of time.

Though not evident, the research design is highly problematic because there were no indicators of student academic performance or cognitive ability when the 95 students were at the *end* of their kindergarten year. Outcome measures from the MAT-6 were initially observed at the end of first Grade, not kindergarten. No measure of student cognitive ability at the time of retention existed because the Otis-Lennon Test was not administered until the end of first Grade. Although the authors presumed the students were of comparable ability at the end of first Grade, the sample sizes of each group were probably not large enough to detect a statistically significant difference. Calculating effect sizes by dividing the difference between group means and the population standard deviation of 16 for the Student Ability Test, revealed that the retained kindergartners were, respectively, -0.50, -0.31, and -0.45 standard deviations behind the (a) transitioned, (b) recommended for retention/transition but promoted anyway, and (c) promoted pupils. The magnitudes of these standardised differences suggest that the retained kindergartners were of lower cognitive ability than students in the other three groups. The absence of an indicator of initial ability or an earlier measure of academic achievement as a statistical control nullifies the authors’ conclusions.

The retained and socially promoted Wyoming elementary school students were more carefully matched on academic ability in Ferguson’s (1991) study. Forty six kindergartners were required to attend a one-year transition first Grade while 20 pupils who had also been recommended for the transition readiness program were placed in first Grade, due to parental wishes or limited class size in the transition room. Students were matched on gender, age, and free/reduced lunch participation. Student development was measured by the Gesell School Readiness Test in the spring of the kindergarten year. The readiness scores of the two groups were comparable. Because no statistically significant mean differences between the promoted and retained pupils at the end of second Grade were found, the author concluded there was no academic advantage to placing students in a transition grade instead of promoting them to the next grade. Although some measurement specialists have questioned the validity of using the Gesell Tests to measure

cognitive ability, the retained and control Wyoming students were probably of similar cognitive ability. Hence, the quality of controls is classified as adequate.¹²

Using the same data set, Ferguson reanalysed the Wyoming students when they were in fourth Grade and again found no statistically significant difference between the transition and promoted pupils (Ferguson and Mueller-Strieb, 1996). The authors concluded that making students repeat a grade did not increase later academic performance. Such an interpretation should be qualified because the sample sizes of 33 transitioned students and 14 promoted pupils were even smaller than in the previous analyses of second Graders, which further reduces the likelihood of finding a statistically significant difference. Unlike the prior study, the authors did not indicate if the transitioned and promoted students had similar Gesell Readiness scores. Given the unknown degree of cognitive comparability between the retained and promoted groups, this article is listed as having an inadequate research design. The authors could have strengthened their conclusions by using the initial Gesell score as a covariate to predict the fourth Grade achievement scores.

The remaining retention studies shown in Table 2 were able to obtain initial measures of student academic performance. However, the research designs in Phelps, Dowdell, Rizzo, Ehrlich, and Wilczenski (1992) and McCombs-Thomas et al. (1992) were classified as inadequate because the authors made no effort to control for existing differences between promoted and nonpromoted students at the time of retention. Phelps et al. (1992) attempted to equalise initial differences by matching on individual characteristics across three student groups examined in a blue-collar suburban Buffalo, New York school district. The first set of students, classified by teachers as “slow learners” had been placed in a pre-grade transition room instead of being promoted directly to the next grade. Of the 22 pupils in this group, 16 were placed in a pre-first grade transition room; five students were required to attend a pre-second grade transition room; one student was placed in a pre-third grade room. Another 24 students in grades two through four were viewed by their teachers as developmentally immature. These pupils were required to repeat their present grade. A control group of 24 students were matched on gender, free lunch status, IQ, and current grade placement with the transitioned and retained pupils. The students in the control group were current classmates of the transitioned and retained children who had never been retained. Outcome measures were obtained five to ten years after retention. There were two general findings. First, there were no statistically significant differences in reading scores across the three groups. Second, the control students had significantly higher mathematics scores than the transitioned and retained students. The authors concluded: “Results suggested that neither retention nor transition placement resolved the academic deficiencies of these children” (p.112).

However, this interpretation is questionable because the authors did not adequately control for initial differences in academic achievement. Reading and mathematics scores from the California Achievement Test (CAT) were obtained at the end of first Grade. Instead of reporting ability scores at the end of kindergarten when students were retained, the pretest scores for the 16 transitioned kindergartners (73 percent of the group) were based on their end-of-year pre-first transition grade, one year after their placement. The absence of end-of-year kindergarten test results limits the ability to assess comparability across the three groups. Given that this was a retrospective study, the CAT scores reported five to ten years after being transitioned or retained were obtained from school records, when the students were in Grades seven through nine. While the IQ scores of the three groups were similar, the grade level of initial measurement was not reported. The study does not indicate if IQ scores were obtained before or after retention. A more meaningful summary of the effect of the various educational practices could have been obtained

¹² A later analysis of these data (Ferguson and Mueller-Strieb 1996) revealed that other measures of early development were available (e.g., the Brigance K and 1 Screening Test as well as the Metropolitan Readiness Test). It would have been useful to ascertain if the students placed in the transition room and their promoted peers differed on these measures which may be better measures of student ability than the Gesell School Readiness Tests.

had the authors performed an analysis of covariance using the end of first Grade CAT scores as a covariate to estimate group differences in the later grades.

The degree to which retained and promoted students were similar at the time of retention was also uncertain in analyses based on a small number of pupils in a rural school district (McCombs-Thomas et al., 1992). Students who were retained in kindergarten or first Grade ($n=31$) were matched on race, gender, and grade point average, at the time of retention, with 31 nonretained students. Student grade point averages when the children were in Grades two through five were the outcome measures. No statistically significant differences in grade point averages were observed in any of the grades. However, separate analyses by race revealed the retained white students had significantly lower GPAs in third and fifth Grade than their promoted white counterparts. The authors did not specify if the initial grades were the same between the retained and nonretained white elementary pupils at the time of retention. McCombs-Thomas et al. (p.347) even acknowledged that the two groups of students may not have been comparable because teachers described the retained pupils as progressing much more slowly in their coursework than the promoted students. The two groups had similar initial GPAs because teachers graded on the work students had completed; however, the retained students were required to repeat either kindergarten or first Grade because they could not keep up with their promoted classmates. The authors then ignored the problem of comparability by stating that the students could be matched only on these limited numbers of variables. They imply that inability to obtain similar students for comparative purposes should be overlooked when assessing the impact of educational programs.

Nine of the 18 studies cited in Jimerson's review met the criteria required to be classified as having adequate controls, although various dimensions of the studies were still problematic. A sample of 53 retained Marin County, California kindergartners were matched with 53 kindergarten same-grade peers who had not been required to repeat the grade (Mantzicopoulos and Morrison, 1992). Pupils were matched on age, socioeconomic status, and academic achievement, calculated from the Stanford Achievement Test (SAT) or the California Test of Basic Skills (CTBS). Further, a screening instrument was used to determine if the pupils were at-risk for reading failure. The two groups initially differed by 0.2 of a standard deviation in achievement scores. The probability of a retained student being at-risk of failing reading was 0.71 whereas the likelihood of have reading difficulty among the promoted was 0.51. After repeating kindergarten, the retained students' second set of kindergarten reading and mathematics achievement scores were significantly higher than the initial kindergarten scores of the promoted pupils. However, because there were no statistically significant lasting gains from the first and second Grade comparisons, the authors presumed that making students repeat a grade was not beneficial.

Two issues make interpretation of the findings difficult; the first pertaining to measurement procedures. The retained kindergartners were pooled from two different school districts in two different academic years. Given that one district used the Stanford Achievement Test to measure academic achievement and the other district the California Test of Basic Skills, the authors converted the initial responses from the two different tests to standard scores in an attempt to make them comparable. It is unclear as to how the z scores were computed. Combining z scores based on national norms from two different tests is problematic because the norms were not based on the same populations. Only if all the retained and promoted students took both the SAT and CTBS would z scores be appropriate to create a combined index of academic achievement, but these children took different exams. A more problematic issue is that the authors ignored the significantly higher observed levels of immaturity among the retained students (Mantzicopoulos and Morrison, 1992, p.193). The promoted and retained students were not comparably matched in abilities that could affect academic achievement. Had the study utilised initial levels of academic performance and behavioural characteristics as covariates, the results could have been more easily

interpreted because plausible rival explanations of earlier group differences would have been eliminated.

In a later article Mantzicopoulos (1997) attempted to gauge the impact of grade retention among the students with high inattention scores. Twenty five children with high inattention (that is, low maturity) scores were contrasted with 15 control students from the initial study who also had high inattention ratings (Mantzicopoulos and Morrison, 1992). Math and reading achievement *z* scores derived from the two different tests from the earlier study were again the outcome measures of achievement. Children's attention problem scores were used as a covariate. Findings based on both same-grade and same-age comparisons indicated that the retained students outperformed the nonretained pupils on the mathematics test, but no advantage occurred on the reading test. As was the case with the earlier study, the initial achievement measures were not used as covariates. The same-grade comparisons were made only in first and second Grades. A more appropriate time frame would have been to use outcome scores as was done in the first paper based on the total number of measurement periods, that is, the contrast between retained and promoted children at the end of the second year of kindergarten, first Grade, and then second Grade. The initial paper revealed that the positive impact of holding children back occurred during the year of retention, not the year after. Mantzicopoulos's findings may be biased because initial scores from the end of the first year of kindergarten were ignored.

A similar analysis was conducted on Tennessee students. Rust and Wallace (1993) matched 60 students, who had either been retained in kindergarten or placed in a transition room, with 60 low-achieving pupils who were promoted to first Grade. Students were matched on race, gender, free lunch status, and classroom grades prior to retention. Outcome measures were student grade point averages and national standardised tests (that is, the California Achievement Test and various forms of the Stanford Achievement Tests). The authors did not explicitly state if same-age or same-grade comparisons were made, however, readers were given the impression that nonpromoted and promoted pupils were evaluated when the promoted pupils were one year ahead of the retainees (that is, same-age comparisons). Although both groups of students had similar grades at the end of the kindergarten year, the retained/transitioned pupils scored significantly lower on the kindergarten achievement test. The average nationally-normed test score of those students who were retained at the end of the school year was 0.27 standard deviations below that of the promoted students. At the end of the next year pupils were retested, after the retained/transitioned students were held back and the nonretained students were at the end of the first Grade year. The retained students obtained significantly higher scores on the national standardised examination than the promoted; the effect size was 0.27 in favour of those held back a year. However, no significant differences in test scores were observed over the next two years. Classroom grades for the retained/transitioned children were also higher in the second and third years of the study, but the differences with the promoted children were not statistically significant. The authors argued against holding students back a year because "This study found weak evidence that retention may benefit children" (Rust and Wallace, 1993, p.165). Given that comparable nonretained low-achieving students were performing successfully in the next grades, the authors did not recommend making students repeat a grade. As occurred in the previous studies, the authors did not statistically control for initial differences in test performance. Using the kindergarten test results as a covariate to help control for initial differences in academic ability between the promoted and nonpromoted children may have resulted in a more accurate assessment of the impact of grade retention.

Jimerson's review of retention studies also cites his research which tracks the academic success of a small number of Minnesota children through their teen and early adult years (Jimerson, Carlson, Rotert, Egeland, and Sroufe, 1997). Thirty two students who had been retained between kindergarten and third Grade were initially matched with 50 equally low-achieving children in the same grades who had been promoted. Nonretained students who fell in the bottom quartile of the

Peabody Individual Achievement Test (PIAT) in first, second, or third Grade were selected for comparison purposes. Teacher ratings of kindergarten students were used to obtain a comparison group for the earliest retainees because the PIAT was not administered at the end of kindergarten. Children required to repeat kindergarten were given the PIAT at the end of their second year in the grade. Retained students and their matched counterparts were found to have statistically similar PIAT scores as well as similar levels of intelligence, as measured by the Wechsler Preschool and Primary Scales of Intelligence (WPPSI). Relative to the low-achieving promoted control group, the retained students evidenced greater behavioural and emotional-health problems when in kindergarten.

The first set of analyses assessed the short-term effects of holding students back in grade by examining academic outcome indicators after the year of retention. Measures of academic performance among the kindergarten retainees, however, were taken at the completion of first grade. Given that only nine retained kindergarteners were contrasted with 15 matched pupils, it is to be expected that no statistically significant differences existed between the two groups. In addition to the small numbers of students for comparative purposes, no data existed to measure the degree of academic achievement when both the retained and promoted students would have been completing kindergarten. The retained pupils could have had significantly lower levels of achievement at the end of kindergarten than their same-grade classmates. A stronger case for the ineffectiveness of kindergarten retention could have been made by statistically controlling for the Wechsler IQ measure (WPPSI), which was taken when the students were age five, prior to kindergarten retention. Although the means of the two groups were not statistically different, the retained children were 0.2 of a standard deviation behind the low-achieving promoted students when tested in kindergarten.

A better statistical adjustment was possible when examining the impact of first or second Grade retention because the PIAT achievement score at the time of retention was entered as a covariate to predict later academic achievement. The retained students significantly outperformed their low-achieving promoted classmates in mathematics, but not in reading. Again, small samples reduced the power of statistical tests as only 16 retained pupils were matched with 28 promoted students. Additional comparisons between the retained and matched students were made at the end of sixth Grade and when the students were age 16. At both times, no significant mean differences in academic achievement emerged.

Jimerson (1999) also reported on the academic achievement when study participants were in Grade eleven and later. The number of study participants with complete data decreased considerably over time. Only 20 students who had been retained before fourth Grade were available to be compared with 23 matched controls. Because many of the student's family, social, and earlier levels of academic achievement were presumed to be similar, Jimerson made no attempt to control for these groups differences. He found that, when compared to their promoted low-achieving counterparts, the retained students had significantly lower levels of academic adjustment when in 11th Grade, were more likely to have dropped out of high school by age 19, and have a lower probability of completing high school by age 20. Jimerson concluded retention led to "poorer" academic outcomes.

The first study (Jimerson et al., 1997) attempted to statistically control for baseline measures of academic performance. Similar statistical controls would have resulted in greater credibility to his conclusions in the second study (Jimerson, 1999). The earlier study indicated that, when in the elementary grades, students were dissimilar on important characteristics that could affect later academic achievement. For example, the measure of student intelligence (Wechsler Intelligence Scale for Children-Revised), taken when the children were in third Grade, was found to be significantly different between the retained and promoted students. The retainees were 0.37 of a standard deviation below the mean score of their matched counterparts prior to high school

(Jimerson, 1999, p.257). Another indication that the promoted and nonpromoted students were dissimilar was that the retained students had missed a significantly greater percentage of school days than the other low-achieving children (p.257). State laws often mandate that students repeat a grade if they are absent a certain number of days. Although Jimerson noted that various reasons can be used to retain students, he made no attempt to determine if differences existed in the intelligence or academic achievement of students who were retained, probably because of absences, with those who were promoted. For these reasons, Jimerson's (1999) later article is labelled inadequate.

These initial differences between the two groups of students raise the question as to whether the later achievement differences between the retainees and the promoted can actually be attributed to making students repeat a grade in elementary school. The high school measure of academic adjustment Jimerson created is problematic because it is a composite index of three separate indicators: high school achievement (grade point average and ratio of high school credits to number of years in high school), behavioural problems, and attendance. However, the retained and socially promoted peers differed significantly in the elementary grades on the behavioural problems and attendance measures. Even with small sample sizes, it would have been appropriate to incorporate the earlier elementary school indicators of these behaviours in separate regression equations so as to help rule out the possibility that differences observed in high school were not due to initial behavioural differences between the socially promoted and retained students. Similarly, the PIAT first grade measure of academic achievement and the third grade WISC-R could be used as covariates in the analyses to predict later academic achievement. Like many other analyses of the impact of grade retention on student academic performance, both the Jimerson studies could have better controlled for initial differences between the promoted and nonpromoted children.

One of the few other analyses listed in Jimerson's (2001) review which utilised regression techniques to control for initial differences between retained and promoted pupils followed the academic achievement of inner city African American children attending Chicago public schools (Reynolds, 1992). The outcome measures were Iowa Test of Basic Skills (ITBS) reading and mathematics scores, along with teacher evaluations of the students' academic competence. Two hundred and thirty one elementary school students who had been required to repeat any grade between kindergarten and third Grade were contrasted with 1000 "regular" children who had never been retained. More important was the availability of a control group consisting of 200 similar low-achieving students who had not been retained. In addition to establishing a control group of students with matched ability, social, and, psychological characteristics similar to the retained children, Reynolds statistically adjusted for possible differences between the two low-achieving groups. For example, Grade one ITBS scores were included as covariates to control for possible differences in academic ability. Same-age comparisons were made in which the promoted pupils were in Grade four while the retained students were still in Grade three. Consistent with findings from other studies, the academic achievement of retained pupils significantly lagged behind the normal students who had never been required to repeat a grade. More importantly the retained students also exhibited lower scores in reading and mathematics than the low-achieving control group of promoted children. The large negative effect sizes for the retained students led Reynolds (1992, p.117) to conclude that the effects of grade retention were negative and harmful or, at best, negligible.

Jimerson also cited two other studies based on the same data. Reynolds reported the results from an almost identical analysis, except only reading scores were the outcome measure (Reynolds and Bezruczko, 1993). Whereas the earlier 1992 study examined ITBS grade-equivalent scores, the outcome measures in the 1993 article were transformed into logit values derived from a one-parameter item response theory method. Unlike the previous study, the control group consisted of all students who had never been required to repeat a grade. The control group of comparable low-

achieving students described in the 1992 analyses was not utilised. Grade retention was again found to have a negative net effect on reading performance based on the transformed ITBS reading scores. Academic achievement levels of the same retained and promoted Chicago elementary school students were also assessed later when the adolescents were age 14 (McCoy and Reynolds, 1999). Both same-age and same-grade comparisons in this last article revealed students who had been retained at some point between Grades one and seven significantly underperformed in both reading and mathematics when compared to the continually promoted adolescents.

The Reynolds analyses are superior in several ways to most other retention studies. The large numbers of retained and promoted students enable detection of significant differences unlikely to be found in studies with small samples. A major positive feature is that, rather than only relying on a matching procedure, Reynolds was able to statistically control for many determinants of academic achievement prior to or near the time of retention. Most important is the availability of similar low-achieving pupils reported in the first study (Reynolds, 1992) which allowed for the creation of a control group more comparable to the retained pupils. All of these features result in a more accurate assessment of the impact of grade retention on academic performance. Nonetheless, the Reynolds papers possess features that raise questions about the certainty of the conclusions.

The higher academic achievement of the promoted pupils is not unexpected, even after statistically controlling for initial differences in family background, cognitive readiness, and other psychological traits, because the total promoted group of pupils ($n=1000$) evidenced significantly higher baseline scores. Jackson (1975) argued that such a “comparison is biased toward indicating that grade promotion has more benefits than grade retention because it compares retained students who usually are not having as severe difficulties, as evidenced by the fact that they have not been retained in grade” (p.619). Campbell and Kenny (1999) point out that if the treated group (for example, the retained students) has lower scores than the control group (for example, promoted pupils), an analysis of covariance will still likely under adjust for initial differences because of the “failure to measure and control for the variable that is used to assign persons to treatment groups” (Campbell and Kenny, 1999, pp.75-76). Within the context of the effects of grade retention, the inability to properly specify those variables that actually caused the student to be retained yields coefficients that imply holding students back a grade is more negative than is likely to be the case. A related shortcoming is the absence of the comparable 200 low-achieving children in the control group reported in the initial study (Reynolds, 1992) which was not utilised in either of the latter two papers (that is, Reynolds and Bezruczko, 1993; McCoy and Reynolds, 1999). Given the greater equivalency in ability measures of these matched students with the retained pupils, findings indicating higher academic achievement of these students compared to the retained would have further strengthened the argument against retention.

Closer examination of the reported differences in academic performance between the retained and comparable promoted students in the initial article (Reynolds, 1992) raises additional concerns about the findings. Reynolds analysed ITBS scores which had been converted to grade-equivalent scores. Although Reynolds (1992, pp.118-119) noted that grade level scores were sometimes misinterpreted, he assumed the psychometric properties of grade level scores were adequate for comparative purposes. However, he overlooks the criticism that grade level scores do not allow adequate measurement of change over time between groups.¹³

¹³ Reynolds (1992, p. 118) attempted to justify the use of grade-equivalent scores by noting that their correlation with logit values is 0.98; however, readers are not informed about the nature of the correlations between grade-equivalent and logit scores over time for the three groups of students examined. Equal-interval logit scores appear in his later two articles (Reynolds and Bezruczko, 1993; McCoy and Reynolds 1989).

The type of group comparison Reynolds (1992) used may be unclear to the casual reader. For simplicity Reynolds used the term *same-grade* to describe the nature of the comparison between the retained and promoted students (p.104). However, the term *same-grade* is commonly used to denote a comparison when pupils are in the same grade. The appropriate definition for the comparisons Reynolds performed is *same-age* because outcome measures were based on scores measured in the same year when students were in different grades (in this study Grade three for the retained pupils and fourth Grade for the promoted students). Because nonretained students cover an additional year of instruction in a later grade, same-age comparisons usually result in the promoted students demonstrating higher levels of achievement than the retained. Having been exposed to a more advanced curriculum for about eight months would help explain why the ITBS reading and mathematics grade-equivalent scores of the matched promoted group were about 0.8 of a year ahead of the retained students; that is when the retained students were in third Grade and their promoted counterparts were in Grade four.

Had Reynolds (1992) compared the academic performance of the 200 pupils in the matched promoted group students when both sets of students were in third Grade, (instead of different grades the same year), the negative results of making students repeat a grade may not have been as pronounced. The matched promoted students would not have had an additional year of exposure to the fourth Grade curriculum which probably raised their grade-equivalent test scores. Although not a same-grade comparison as is traditionally defined in retention research, Reynolds (1992, p.113) compared grade-equivalent test scores for the retained students when they were in Grade three with those of "matched-grade peers." These third Grade students were likely to be a year younger than the retained students because these matched-grade peers had never been required to repeat a grade. Although the effect size for reading was -0.12 and 0.07 for mathematics, the third Grade test results did not significantly differ. Reynolds (1992, p.113) therefore concluded that "retention is unrelated to academic achievement" because he assumed the retained students should have had higher levels of academic achievement since they were a year older. Such an assumption seems questionable because the most advanced curriculum the retained students had been introduced to was the same as that of their third Grade classmates. Although the effect sizes between the two groups of children were negligible and statistically insignificant, it is not unreasonable to speculate that the initial first Grade ITBS scores of the retained third Graders were significantly below those of their nonretained third Grade classmates, otherwise they too should have been held back a year. Were the matched-grade peers to have higher scores in first Grade than the retained students, one could reasonably argue that making the low-performing pupils repeat a grade helped bring them up to the level of their classmates who were continually promoted to third Grade. Reynolds provides no information in his article to assess the reasonableness of this hypothesis.

Only two of the studies examined in the Jimerson (2001) meta-analysis concluded that requiring low-performing students to repeat a grade improved academic achievement. Pierson and Connell (1992) compared the academic performance of 74 upstate New York students retained in grades one through four (most were required to repeat first Grade) with 69 promoted same-grade peers matched on Otis-Lennon Mental Ability IQ scores, sex, and grade when both groups of students were in their current grade. An additional comparison group consisted of 35 students who had been placed in the next grade, that is, the pupils should not have been promoted because of grades or the teacher's recommendation. These children were matched with the retained pupils on the basis of similar grade point average, sex, and grade. Grade point averages were comparable between the retained and socially promoted pupils at the time of retention (Pierson and Connell, 1992, p.303). The overall outcome measure of academic performance was based on the (a) mean value of final marks for all academic subjects averaged from Grades three through six and (b) the mean reading and mathematics results from national standardised percentiles from Grades two

through six (that is, the grades after retention). These two indicators of academic performance were averaged to measure overall achievement.

The retained pupils achieved academic outcome scores similar to their peers with comparable IQs. More interesting was the finding that the retained students performed significantly better on the global indicator of academic achievement than the socially promoted students. The mean academic achievement score for the retained students was 0.56 standard deviations above that of the socially promoted. Consequently, the authors argued that, while not a “cure-all” for below grade-level academic performance, “the findings support the use of retention as a potentially effective remediation for academic difficulty in the early elementary grades” (Pierson and Connell 1992, p.306). A limitation of the study, however, is that no annual academic achievement data were presented, only the aggregate of grades and standardised test scores over the period of study. It is not possible to determine whether the positive impact of retention occurred only after the year of being held back, or whether the difference in achievement scores persisted over time. A unique aspect of this study is that no other articles combined annual performance indicators into a single score; instead, other studies reported yearly achievement measures.

The second study concluding in favour of retention was a prospective study of a random sample of 800 elementary students in the Baltimore City Public Schools (Alexander et al., 1994). Study participants were followed from the fall of 1982 when the students were entering first Grade through the spring of 1990. Over 300 students had been required to repeat a grade at least once. Given the many different kinds of outcome measures analysed, it is difficult to determine the exact number of students who were retained or promoted with usable data. Academic achievement was measured by use of the California Achievement Tests (both reading and mathematics) and course grades in reading and mathematics. There were 242 students who repeated either first, second, or third Grade. The major control group of interest was 106 never-retained poor-performing children, although comparisons were also made with regularly promoted students. Students were examined by grade of retention. The actual number of students analysed varied because pairwise deletion of data was used to maximise the number of observations. After following students from first Grade through eighth Grade, the authors concluded that making students repeat a grade helped boost their academic ability so that they performed at levels closer to those of the regularly promoted students, although the nonpromoted children never caught up with those students who experienced no academic difficulties.

Before evaluating Jimerson’s overall conclusion regarding the meaning of all these studies reviewed, it is worthwhile to point out several types of errors which occurred in his meta-analysis. One type of error Dunkin (1996) identified is “listing different reports from the same project as providing additional confirmation of the same finding” (p.91). This “double counting” appeared several times in the Jimerson (2001) meta-analysis. Two studies were based on a small sample of Wyoming elementary school children (Ferguson, 1991; Ferguson and Mueller-Streib, 1996). Students sampled from Marin County appeared in two papers (Mantzicopoulos and Morrison, 1992; Mantzicopoulos, 1997). Jimerson also cited his two analyses which used participants from the same Minnesota sample (Jimerson et al., 1997; Jimerson, 1999). Similarly, findings from the three papers by Reynolds and his colleagues (Reynolds, 1992; Reynolds and Bezruczko, 1993; McCoy and Reynolds, 1999), which utilised the same sample of low-income African American children from Chicago, were counted as if they were from different data sets. In all of these articles, the authors concluded that grade retention diminished student academic achievement. Following the same students over time will likely yield the same negative comparisons between retained and promoted students favouring the latter. The result of treating identical data sets as if they were independent results in an inflated number of negative outcomes.

Syntheses of literature are sometimes in error because an author incorrectly describes the methodology or context of a study. Dunkin (1996, p.90) refers to this kind of incorrect statement

as “erroneous detailing.” Jimerson (2001) made this error several times in his meta-analysis when listing the characteristics of students in the control groups. The first instance occurred when describing the Hagborg et al. (1991) study. Jimerson listed the article as controlling for academic ability because the authors presumed students in the same high school classes were of equal ability. In actuality, there were no prior indicators of student ability. Jimerson committed another “detailing error” when describing the Dennebaum and Kullberg (1994) study because the measure of ability used to match students was not made at the time of retention but after. His description of the McCombs-Thomas et al. (1992) analyses was partially incorrect because, contrary to the information listed in Jimerson’s Table 1, IQ measures were not mentioned as a basis of matching students. The authors of these three studies concluded that retention was ineffective. However, these errors in study details may incorrectly lead readers into thinking the retained and promoted students were more similar than was the case. Another detailing error arose when describing the comparison groups in the Pierson and Connell (1992) study which concluded in favour of retention. Jimerson correctly described one comparison group (n=69) which did not have data indicting initial academic ability. However, Jimerson ignored a second comparison group (n=35) of socially promoted children matched on similar grade point averages with the retained students. These socially promoted pupils evidenced significantly lower grade point averages than the retained at the end of the study period. Given the large number of articles and variables reviewed, it is quite easy to incorrectly list the detailed characteristics of an article. Nonetheless, these kinds of errors may shade the overall conclusions to be made when aggregating the findings from individual studies.

Although Jimerson may not have committed a detailing error per se when interpreting the study of Baltimore elementary students, which concluded in favour of retention, a different interpretation of the findings is possible. The Alexander et al. (1994) Baltimore study is particularly important in Jimerson’s review because 47 per cent of the negative findings reported in his meta-analysis were derived from this analysis. Jimerson summarised both the same-age and same-grade comparisons Alexander and his colleagues calculated. A count of the same-age comparisons resulted in 23 significant negative effects, and seven contrasts which did not differ. These comparisons were based on mean differences between the retained and a group of poor-performing children, after adjusting for initial social-demographic differences and an indicator of student test performance prior to retention between the two groups. With respect to same-grade differences, Jimerson reported a total of 21 statistically significant negative effects, 28 nonsignificant coefficients, and only one significant positive effect for grade retention. These comparisons do not include the results measured at the end of the retention year.¹⁴ In addition, Jimerson ignored a set of more refined comparisons which also controlled for students who were retained more than once and students classified in special education. Under these different circumstances the summary of findings changed considerably. A recount of the same-grade comparisons pertaining to test scores and classroom grades revealed only eight significant negative effects and 49 insignificant differences. Three contrasts favoured the retained students. A reexamination of the Alexander et al. (1994) study suggests that making students repeat a grade is more positive than Jimerson reported. This conclusion, however, is based on a different set of comparative criteria which controls for more factors. Rather than count the number of negative, insignificant, or positive results from specific contrasts between the retained and poor-performing control group, Alexander and his colleagues argue that it is more important to examine the overall

¹⁴ Professor S. Jimerson (personal communication, June 22, 2004) was kind enough to provide me information describing how he derived the numbers summarizing the number of statistically insignificant, negative, and positive mean differences between the retained and poor-performing promoted students in the Alexander et al. (1994) study. The aggregate counts in Jimerson’s (2001, p. 426) summary were derived from the information Alexander and his colleagues presented in Chapters 5 and 6 of their book. The number of insignificant comparisons is actually 38 instead of the 28 shown in Jimerson’s Table 1.

pattern of findings. The more worthwhile question from their perspective is whether retention helped raise the academic performance of children required to repeat a grade. The 12 same-grade comparisons presented in Alexander et al. (1994) which control for double retention and special education characteristics reveals that prior to being held back, the retainees were significantly below the comparisons group. But after repeating the grade, the test scores and grades of the retained students were equal to those of the promoted poor-performing children. Consequently, Alexander et al. (1994) contend their findings demonstrate that retention can help enhance academic achievement because “for retainees just to be holding their own may be an accomplishment” (p.22).

Sixteen of the 18 studies listed in Jimerson’s meta-analysis did not favour making students repeat a grade to improve their educational performance. As was also observed in the Holmes (1989) review, authors of individual articles overlooked findings which contradicted their overall conclusions. Partial results from four of the studies in the Jimerson meta-analysis could be interpreted to suggest that grade retention is an effective strategy to help raise student achievement. Because no statistically significant differences were found between retained and promoted students, various authors concluded holding students back a year in school was ineffective. Had the authors computed effect sizes between the two groups, they may have changed their conclusions. For example, the investigation of Washington state students (Johnson et al., 1990) failed to mention that fourth-Grade total math and math calculation scores were higher among the retained. If one computes effect sizes for these two subscales from the Metropolitan Achievement Test, the respective values are 0.44 and 0.42, which imply that retention helped increase mathematics performance. Similarly, Ferguson’s (1991) analysis of Wyoming pupils showed the mean values from the SRA Achievement Test taken in second Grade of children put in the readiness transition room were numerically larger than those of the promoted kindergartners. Computing effect sizes for the mean differences yielded values in favour of the retained pupils. The effect sizes for each tested area were 1.17 (language), 0.86 (math), 0.28 (reading), and 0.88 (total score). These findings imply that placement in the transition first Grade was helpful, particularly in three of the four areas tested. Had Jimerson et al. (1997) calculated effect sizes between retained and nonretained students for certain outcome measures, they would have found that the children who repeated a grade outperformed those who were promoted. The effect size between the two groups for the total score of the PIAT at the end of first Grade was 0.38 standard deviation units in favour of those retained in kindergarten. Among the first Grade and second Grade retainees, the effect sizes for the PIAT mathematics test were substantial for both the unadjusted means (1.21) and the adjusted means (1.07). With respect to the PIAT total score, the retained first and second Graders outperformed the promoted students by 0.74 standard deviations on the observed means and 0.64 standard deviations for the adjusted means. If the authors of these three studies assumed effect sizes of 0.20 or higher indicated “educationally meaningful” differences, as Reynolds (1992, p.107) suggested, they could have concluded that holding low-performing students back a year provided an educational benefit. None of these authors calculated effect sizes because there were no statistically significant differences between the retained and promoted students. However, the number of observations in each comparison group was so small that even substantively large calculated effect sizes were deemed insignificant. Group means were not statistically significant because of the small sample sizes with insufficient power to reject null hypotheses, even when the levels of academic achievement were substantially different in standard deviation units.

The magnitudes of the effect sizes suggest that the authors’ conclusions regarding their findings may be based more on their subjective interpretations rather than the specific results from the data. Does an insignificant difference between retained and promoted students indicate that retention is effective? The interpretation partially depends on what authors assume is needed as proof of effectiveness. The study of New York students by Phelps et al. (1992) revealed that the

pre-retention reading and mathematics means of the transitioned and retained students were significantly lower than that of the control group. By the end of the period under investigation, however, mean reading values were comparable across the groups. This similarity could be interpreted to indicate that making students repeat a grade helped reduce the initial differences among the three groups. That is, the lower-performing pupils who were held back a year caught up with their promoted classmates. But Phelps and her colleagues assumed that the scores of the retained students should surpass that of the nonretained children. The absence of standard deviations in this study prevents readers from calculating effect sizes between the transitioned, retained, and matched control group. However, an examination of the means shows that the students placed in the transition room raised their reading scores by nine percent while the retained students increased their reading scores by over seven percent. Conversely the reading scores of the promoted control group decreased by almost five percent. Regardless of the authors' conclusion, these findings reveal that holding students back a year at least enables them to catch up with their promoted classmates and not fall further behind.

As has been stressed throughout this discussion, the inability to control for initial levels of academic achievement between retained and promoted groups can lead to only ambiguous conclusions about the impact of holding students back. Even though several of the authors recognised this limitation of their analyses, they nonetheless concluded that low-performing students should not be required to repeat a grade. The Phelps et al. (1992) study aptly illustrates this type of reasoning. The authors' acknowledge that (a) only randomisation studies can prove causation and (b) matching on available variable can not control for all sources of variability (Phelps et al., 1992, pp.121-122). But they made no attempt to statistically control for initial differences between retained and promoted students and negatively assessed the impact of transition room placement and grade retention.

Meisels and Liaw (1993) reached a similar negative conclusion, although they recognised the difficulty in interpreting their findings: "it is possible that the retained students who showed less optimal academic performance in eighth Grade were academically less able or had problems that predated retention" (p.75). Nonetheless, the authors rejected this alternative interpretation because (a) their findings were similar to those summarised in the Holmes (1989) meta-analysis and (b) they analysed a large national representative sample of students. As was previously mentioned, a detailed examination of the published retention studies in the Holmes (1989) summary does not yield a consistent pattern of negative effects because most of the research designs do not incorporate baseline measures of student ability. The fact that most of the retention studies lack an earlier indicator of student academic performance does not justify the Meisels and Liaw (1993) conclusion. Hagborg et al. (1991) likewise concluded that making students repeat a grade is probably ineffective because the retained students did not perform as well as their continually promoted classmates. The lack of earlier data measuring student aptitude, which could indicate how far behind the retainees fell relative to their same-grade peers, does not allow Hagborg and his coauthors to infer that retention was not helpful. Rather than recognise the limited nature of their study, the authors relied on the questionable conclusion of an earlier meta-analysis to support their position: "However, given the doubtful benefits of retention (Holmes and Matthews, 1984), it is possible that the educational needs of retained students were not adequately addressed, and they were left behind their classmates" (Hagborg et al., 1991, p.315). By making no effort to evaluate the validity of many of the conclusions offered in the reviewed studies, Jimerson (2001) committed the "Nonrecognition of faulty author conclusions" discussed by Dunkin (1996, p.91).

Overlooked and More Recent Research

Before summarising the Holmes and Jimerson meta-analyses, it is worthwhile to examine several recent published articles. A summary of the research characteristics of these studies is presented in Table 3. One investigation analysed the academic achievement of largely Hispanic elementary

children in south central California (Cosden, Zimmer, Reyes, and Gutierrez, 1995). Kindergarten students who had been required to repeat the grade were matched on birth month, gender, ethnicity/home language with an equal number of classmates at the end of first Grade who had never been retained. Low-performing kindergartners who had been “advanced” (that is, placed) in first Grade in spite of observed academic difficulties were also matched using the same variables with an equal number of never-retained classmates. Achievement was measured at the end of the first Grade year by the Stanford Achievement Test for English speaking students and the Aprenda among Spanish speaking children. Preliminary MANOVAs on reading, language, mathematics, and the basic battery indicated that the English speaking retained (n=17) and advanced kindergartners (n=35) had lower scores than their never-retained classmates. Mean first Grade scores on the Aprenda were comparable between the retained Spanish speaking children (n=19) and their classmates, but the reading scores of the Latino students who had been advanced into first Grade were significantly below their matched promoted Spanish speaking first Grade peers. Multiple regression analyses which combined all of the students revealed that only the low-achieving advanced children (both the English and Spanish tested students) had significantly lower scores than their controls. Cosden, et al. (1995) concluded “that neither intervention [retention or advancement] had a positive impact on the students’ achievement” (p.137). However, the authors correctly acknowledged that the study could not assess the extent to which retention or advancement influenced academic results because of uncontrolled factors, such as parental education and quality of relationships with the schools. Like most of the studies cited in the retention meta-analyses, no baseline measures of kindergarten performance were available as control variables. The academic skills of the retained and advanced students were probably below their first Grade classmates even in kindergarten.

Table 3. Characteristics of Additional Retention Studies

Author(s)	Quality of Controls	Outcome Pretest	Ability Pretest	Statistical Control for Outcome Pretest	Statistical Control for Ability Pretest	Type of Comparison	Grade Equivalent Units	Authors' Conclusion Effect of Retention
Cosden et al. (1995) (36-36)	Inadequate	No	No	No	No	Same-Grade	No	Negative
Pomplun (1988) (47-47)	Adequate	Yes	Yes	Yes	Yes	Same-Grade	No	Positive
Southard and May (1996) (66-66-66)	Adequate	Yesa	Yesa	Yes	Yes	Same-Grade	No	Negative ^b
Lorence et al. (2002) (736-28,351)	Adequate	Yesa	No	Yes	No	Same-Grade	No	Positive
Jacob and Lefgren (2004) (8,120-5,018)	Adequate	Yesa	No	Yes	No	Same-Age	No	Positive

^a Retained and promoted students were significantly different on this variable at time of retention.

^b Data indicate that retention had a positive effect on academic achievement.

Findings from two other studies with better controls for initial ability suggest that making students repeat a grade may increase academic achievement. The first study investigated the effect of making students in a west central Florida semirural county school district repeat a grade. Pomplun (1988) examined the change in NCE scores on the Comprehensive Tests of Basic Skills among retained students, borderline pupils who had been placed in the next grade, and regularly promoted students. Although retained and nonretained students were matched on gender, grade, age, self-concept, and level of motivation, the retainees were acknowledged to differ substantially from their promoted peers on unspecified sociological and psychological variables. Twenty-two students were retained in Grades one or two; 15 students were required to repeat either Grade three or four; 10 pupils in either Grades seven or eight were also retained. A comparable number of borderline or regularly promoted students were matched with the retained children in each of

the three broad grade categories. Students were retained in the spring of 1983. All students whether retained or promoted were again tested in the spring of 1984. Retained pupils in the primary and intermediate elementary grades evidenced significant increases in reading, language, and mathematics while the promoted borderline students showed a decrease in achievement during the same year. Among the seventh and eighth Grade students, however, changes in achievement scores were indistinguishable between retained and the borderline placed pupils. Pomplun suggested that retention was more beneficial in the earlier elementary Grades than in middle schools. A limitation of the analyses is that students were followed for only one year.

Another study (Southard and May, 1996) incorporating measures of student ability prior to retention was based on students from three elementary schools in a suburban New York district. Kindergarten teachers rated 66 pupils (from 1982 to 1985) as being unable to perform first Grade work. These students were placed in pre-first grade transition classrooms. Several different comparison groups were analysed. One was a group of 24 students who entered kindergarten at the same time as the transitioned pupils. However, after being promoted to first Grade, these 24 children were required to repeat the grade. Another control group consisted of 66 regularly promoted students who had not been placed in a transition room nor retained in first Grade. Non-retained first Grade classmates of the 66 pre-first grade transitioned students composed the final control group. All students had taken the California Achievement Tests in listening and mathematics at the end of kindergarten. These variables were used as covariates to assess differences in group performance at the end of the first Grade year because both the transitioned and eventual first Grade retainees had significantly lower mean listening and mathematics scores than the regularly promoted pupils. Initial CAT scores were not the primary basis for retention; instead, student behaviour and classroom skills in kindergarten were the sources of placement recommendation. Achievement outcomes were taken from the reading and mathematics sections of the Iowa Tests of Basic Skills when all students were in first, second, fourth, and fifth Grade. Using the kindergarten listening scores as a covariate, the regularly promoted students and year-younger classmates of the transitioned students evidenced significantly higher reading scores than the children who had been placed in the pre-first grade transition room. Average reading and mathematics scores of the transitioned and eventual first Grade repeaters were comparable. However, children in the pre-first grade transition class reported significantly higher math scores than the regularly promoted and younger aged classmates at the end of first Grade. Nonetheless, the authors believed their findings implied that there was little long-term benefit to retention.

The authors acknowledged that “Advocates of pre-first programs might claim that these mixed results between the pre-first and comparison group students were evidence in support of these programs” (Southard and May, 1996, p.139). But the authors reached a negative conclusion because they assumed that the transitioned students should perform better than their younger classmates and the students who had to repeat first Grade. An alternative interpretation is that the transition year helped the academically challenged kindergartners catch up with the continually promoted students who demonstrated higher levels of mathematical ability. Had the authors been able to measure reading aptitude, instead of facility in listening at the end of kindergarten, the reading performance of the transitioned students may have been similar to that of the regularly promoted children in the later grades.

Another study reporting that grade retention is associated with a positive impact on school performance appeared a year after the publication of Jimerson’s meta-analysis. Lorence, Dworkin, Toenjes, and Hill (2002) tracked the academic progress of a cohort of all low-performing third Graders in Texas public schools from 1994 through 1999. The same-grade scores of over 700 students who failed the state’s mandatory reading test (the Texas Assessment of Academic Skills Test – TAAS), and had to repeat third Grade, were compared with the test results of over 27,000 socially promoted third Graders who also failed the TAAS reading examination. The retained third Graders not only caught up with the promoted students (who

actually had higher reading scores than the retained at the end of third Grade), but the retainees statistically surpassed their low-performing promoted counterparts each year after being held back in third Grade. After statistically adjusting for initial performance levels and socio-economic variables, the retainees were outperforming the socially promoted pupils by 0.30 standard deviations at the end of seventh Grade. Although many of the socially promoted third Graders eventually passed the state reading test, the retained students, on average, passed the test a year earlier. The authors found no evidence that grade retention hindered the academic performance of students over the six years of study.

Two economists (Jacob and Lefgren, 2004) recently evaluated the impact of third Grade and sixth Grade retention among children enrolled in the Chicago Public School System during the mid and late 1990s. Using a regression-discontinuity design which statistically controlled for student prior test performance, in addition to social and demographic characteristics of the children and their neighbourhoods, the authors found that retention helped improve the performance of third Graders in both reading and mathematics. However, retention was associated with somewhat lower scores in reading and mathematics among students who had been required to repeat sixth Grade. The authors caution that the estimated effect of grade retention was confounded with high stakes tests in certain grades which may partially account for the variation in scores by retention grade. Same-year tests were the basis of comparison for retained and promoted Chicago students, but minimum passing scores are required for exams in Grades three, six, and eight before being advancing to the next grade. For example, the retained sixth Graders took the seventh Grade ITBS test (a low-stakes test with no consequence for their promotion to eighth Grade) the same year when the promoted sixth Graders were required to take the eighth Grade ITBS test. The retained sixth Graders were probably less motivated to do well that year than the promoted sixth Graders who needed a minimum score on the eighth Grade exam before being promoted to Grade nine. In spite of the difficulty of interpreting the impact of grade retention, the authors did not find that making low performing students repeat a grade adversely affected their later academic achievement.

DISCUSSION

The Holmes (1989) meta-analysis is the source most often cited as demonstrating the futility of making students repeat a grade to improve their academic shortcomings. Since 1989 it has not been possible to read a discussion on grade retention which does not mention, in some form, Holmes summary that “The weight of evidence argues against grade retention” (p.28). Critics of grade retention practices (for example, Dawson, 1998) maintain that “the most valid and best designed studies...clearly supports promoting underachieving students over retaining them” and then cite the Holmes meta-analysis to support their assertion. Jimerson’s (2001) meta-analysis will also be cited as a seminal study which discredits the practice of retention. Both Holmes and Jimerson are to be commended for their efforts to synthesise studies examining grade retention. However, the major conclusion of the present review is that findings pertaining to the effect of grade retention on student academic performance are not unequivocal; the issue of grade retention has not been resolved. Contrary to the conventional wisdom among educational researchers, this review of the extant grade retention literature argues *there is no overwhelming body of scientifically sound evidence demonstrating that making academically challenged students repeat a grade is ineffective or harmful*. This paper’s conclusion derives from a detailed appraisal of specific studies listed in the two most recent comprehensive meta-analyses of grade retention. A meticulous examination of published studies indicates that the overall quality of research is characterised by serious methodological weaknesses. In-depth inspections of the Holmes and Jimerson meta-analyses reveal considerable shortcomings. Both of these summaries exhibit the many kinds of errors identified by Dunkin (1996) which commonly appear in meta-analyses. The vast majority of studies which conclude that retention is an ineffective educational

practice contain so many limitations that inferences from them are highly questionable if not unwarranted.

Nonetheless, critics of retention argue that the quality of research is sufficiently high enough to discount the practice of making students repeat a grade. The most extreme example of this position is illustrated by Reynolds (1992), who maintains that the quality of findings appearing in the Holmes meta-analysis and more recent research is comparable to those from medical studies:

In medicine, treatments that are shown to be ineffective or to have serious unintended effects do not gain approval from governmental bodies and are subsequently discarded or substantially revised to eliminate their undesirable effects. Despite the accumulated evidence to date, however, retention as an educational treatment has not followed such established scientific traditions. (p.118)

A glaring error in this assertion is the presumption that retention studies have attained the same high standards of research quality to which medical researcher must adhere. Unlike individuals who conduct medical studies, educational researchers have not had the ability to randomly assign participants to specific narrow treatment conditions. Randomisation is a powerful tool which can help rule out alternative explanations for causal effects in medical studies. To imply that retention studies use research designs comparable to the randomised experiments which dominate the field of medical research greatly exaggerates the extent to which results from retention research can be generalised. As Jackson (1975, p.624) noted in his overview of retention, the last time randomisation was used to assign students to repeat a grade was over 60 years ago.

It should be evident from this review of both the Holmes (1989) and Jimerson (2001) meta-analyses that grade retention studies in no way approach the precision of medical research. It is inappropriate to aggregate findings from retention analyses based on inadequate controls and then assume the summary calculations will yield valid conclusions. Few studies examined from these meta-analyses meet acceptable criteria required for reasonable inferences. Valid inferences can not be made from the vast majority of studies in the aforementioned meta-analyses because the two groups of students did not have comparable characteristics when the decision to retain was made. Only four of the ten published articles in the Holmes summary would meet conventional criteria for achieving some degree of comparability between retained and promoted students; however, the quality of the two articles reaching negative conclusions about the usefulness of retention is suspect because of uncertainty in the nature of student differences at the time of retention. When compared to the Holmes review, Jimerson's meta-analysis fares somewhat better because 10 of the 18 studies appear to attempt to control for possible differences between retained and nonretained students, although the comparisons in 4 of these 10 studies are questionable. Rather than relying on a matching procedure to help equalise initial differences between retained and nonretained students, as did many of the articles cited in the Holmes summary, more of the authors in Jimerson's survey utilised regression procedures to make statistical adjustments. Southard and May (1996, p.141) specifically noted the importance of incorporating preretention indicators of outcome measures as covariates to obtain a more accurate assessment of making students repeat a grade; they declared their results would have been very different had they not statistically adjusted initial differences between retained and promoted students.

Even if some of the studies utilised better controls, the small sample sizes and the limited representativeness of the students can yield only the weakest of inferences about making students repeat a grade. Five of the 10 better designed studies analysed small numbers of observations and none of the 10 studies were based on national representative samples of the school population. Findings from the overwhelming majority of retention studies do not support the position that grade retention is an inappropriate remediation practice. This summary is consistent with the position of Alexander et al. (1994) who maintain that the strong opinions individuals hold

regarding the impact of grade retention are actually based on weak empirical evidence from poorly designed studies.

In spite of these methodological shortcomings, opponents of making students repeat a grade will likely continue to cite the Holmes and Jimerson meta-analyses as authoritative proof that retention is an unsuitable educational practice. Nonetheless, the current review has identified seven studies which indicate that making students repeat a grade is associated with higher academic performance. The quality of the research designs of these positive studies are at least comparable to those alleged tightly controlled studies cited in Holmes (1989). Moreover, nine other studies in which the authors do not favour retention present findings which suggest that retention may result in some positive academic outcomes. Whether critics of grade retention will accept these positive studies as “demonstrating the effectiveness of retention as an intervention facilitating subsequent academic success” which Jimerson (1999, p.265) and other critics of retention (see also Holmes and Matthews, 1984, p.232) have required proponents of the practice to provide is unlikely.

Alexander et al. (2003, pp.16-20) contend that educational researchers have such strong negative opinions on the subject of grade retention that they are biased against any evidence which contradicts the view that holding students back a year in grade is a bad practice. Partial support for this view is evident in the kinds of criticisms made against studies which favour retention. Tanner and Galis (1997) have also suggested that “Even when the research is refereed and published we find that sometimes the reporting appears to be biased and misleading” (p.110). For example, Shepard, Smith, and Marion (1996) published a very detailed critique of the Alexander et al. (1994) study. Shepard and her colleagues speculated at great length on shortcomings and alternative interpretations of the Baltimore results. They concluded that the findings of Alexander et al. (1994) were flawed because of the manner in which scores were scaled; special education students should have been removed from the analyses; selection effects and regression artifacts resulted in erroneous findings.¹⁵ Likewise, Shepard (2002) is highly critical of the Texas study (Lorence et al., 2002) which indicated that grade retention may have helped improve the academic competencies of low-performing elementary school children. Her major arguments are that: the authors did not adequately control for the special education students; improvement in test scores was largely the result of regression to the mean effects; selection effects yielded biased results; and public school teachers teach only information pertaining to the state’s mandatory accountability test. My point is not that studies in favour of retention are beyond criticism; but that a more rigorous standard of what constitutes “acceptable research” is applied to studies which contradict the prevailing view on grade retention. Studies which conclude that retention is an ineffective remediation practice are not subjected to the higher standards of methodological rigor required for studies which are critical of letting academically challenged students automatically proceed to the next grade level.

The major purpose of this review, as is indicated in the opening paragraphs, is to draw attention to the situation that most of the studies reported in the research literature on retention are not sufficiently sound to support the claim that grade repetition is always wrong. More recent research suggests that grade retention may indeed help improve student learning. If thoughtful readers will make more of an effort to judge the evidence on retention objectively, this review will have served its purpose. Over a decade ago Kaestle (1993) discussed the “awful reputation of education research.” He pointed out that policies and practices were sometimes highly politicised because the research did not allow the educational community to reach a consensus on important issues. The same can be said about educational researchers who disagree on the utility of making students repeat a grade. Although most educational researcher believe that making students repeat a grade is ineffective, this review challenges that position. A detailed examination of retention

¹⁵ See Alexander et al. (2003, pp.265-279) for a detailed rebuttal to the criticisms of Shepard et al. (1996).

studies pertaining to academic achievement indicates there is no cumulative research that yields a firm conclusion on the topic. Instead of lauding research which supports a specific view and criticising research which contradicts a favoured perspective, a better approach is to seek consensus as to what appropriate criteria are needed to determine if retention is worthwhile and then perform the necessary research which will provide answers. Karweit's (1992) earlier attempt to raise such issues has largely been ignored. For example, there seems to be little agreement as to what the goals of retention should be. Must retained students surpass the academic achievement of regularly promoted peers to be successful, or will merely catching up with socially promoted low-performing students met the criterion of effectiveness? Researchers have also differed in their definitions of retention and the kinds of educational practices which occur during the repeated year. These and other issues need to be addressed before judging the usefulness of retention practices.

SUGGESTIONS FOR FUTURE RESEARCH

Several dimensions of retention research must be attended to if a more definitive assessment of the impact of making students repeat a grade is to emerge. One issue is that the impact of retention may vary by grade. To illustrate, Peterson et al. (1987) found that the learning gains obtained among second and third grade retainees persisted over time while the initial improvement observed among the retained first graders declined. Pierson and Connell (1992) reported that students retained from third to sixth grade experienced learning gains compared to matched counterparts. The Baltimore study found that children held back in first grade, or who were eventually placed in special education, did not appear to benefit as much from retention as those students retained in second or third grade. Alexander et al. (2003) suggest that children with the most noticeable learning disabilities are the first to be retained, usually in kindergarten or first grade. Such children are often much farther behind their peers in academic ability than children retained in later grades. Pupils retained in second or third grade were probably closer in academic ability to their promoted classmates. Because they did not have the kinds of problems typical of earlier retainees, the second and third grade repeaters were better able to learn the material during their year of retention. These findings imply that holding children in kindergarten or first grade may not lead to the same kinds of increased test scores observed among students in higher elementary grades. Pomplun's (1988) study of Florida students further reveals that retaining high school students will be less effective than requiring low achieving students in primary and intermediate grades. Poorly performing high school students may be so far behind their promoted classmates that making them repeat a grade will not enable them to complete a degree. Whether differences on the impact of retention across grades results from differences in the abilities of students held back in various grades should be examined in more detail.

A second issue requiring further research is disentangling the effect of retention with the specific instructional practices provided during the retention year. Karweit (1992) listed several kinds of educational practices offered during a repeated grade. Probably the most common is recycling students through the same grade with no additional resources or special assistance. The students simply repeat the same curriculum. This form of retention may not be helpful. A case in point is the Chicago Longitudinal Study Data, the basis of the three Reynolds' papers on grade retention which are often cited as showing that making students repeat a grade is ineffective. Thus far no one has suggested that the negative effects of retention observed among Chicago public school students may result from ineffective educational practices endemic to that specific school system. During the 1980s through the mid-1990s, Chicago was reputed to have one of the worst public school systems in the nation (see Hess, 1995; Vander Weele, 1994). Chicago school teachers may have been less concerned about helping remediate low-performing students than occurred in other school districts across the country. Indeed, Reynolds et al. (1997) describe the nature of retention practices in Chicago as follows: "Once students are retained, however, they usually get no special

help with their schooling. They are often placed in low academic tracks only to repeat the previous year's instruction and ultimately disengage from school" (p.36). Merely repeating the failed grade may not help a student, but school districts more responsive to the needs of academically challenged students should be examined because the context of the school system may also affect student academic performance independent of making students repeat a grade.

Instead of merely recycling low achieving students through the same grade, another retention practice is to require students to repeat a grade, but also provide them additional learning opportunities during the year of retention. Several studies indicate that providing additional assistance and special programs to retained students may be more beneficial than only repeating the same curriculum. The longitudinal study of children in Mesa, Arizona (Peterson et al., 1987) showed that elementary school children whose teachers developed individual educational plans to address the retained students' academic shortcomings maintained higher scores than their socially promoted counterparts. In their analyses of Minnesota children Jimerson et al. (1997) also "noted that many students in the retained group received additional academic services the repeated year" (p.21). The authors speculated that this extra assistance may have partially accounted for the higher math performance of the retained students. Likewise, Lorence et al. (2002, p.44) commented that interviews with Texas teachers and educational administrators revealed that Texas elementary students who were required to repeat a grade often received various forms of additional educational assistance. These examples imply that retention itself is not likely to be an effective remediation strategy when students are only recycled through the same educational program of the failed year. However, when combined with strategies focusing on the unique academic weaknesses of students, retention may help raise the achievement levels of low-performing pupils. These findings contradict the negative view that "the effects of most retention plus remediation approaches are likely to be disappointing" (McCoy and Reynolds, 1999, p.295).

A third kind of retention practice is placing academically challenged children in an alternative program, often a transition classroom, before actually classifying the students as failing. These transitional classes are usually created for kindergartners or children in first grade. However, the specific instructional practices available to children in these transitional classrooms are seldom described. One exception is Leinhardt's (1980) study of low-achieving kindergartners. Although this study has been cited as indicating that extra educational assistance combined with retention is ineffective, the special instruction offered the retained children placed in a transitional room was of much lower quality than provided socially promoted peers. Before making general conclusions about the effectiveness of retention combined with special help, more specific information is required to learn the specific retention policies students are subjected to and how they are implemented. Researchers should attempt to identify the kinds of specific educational activities occurring during the retention year, instead of grouping different kinds of instructional practices as referring to retention in only a generic sense.

Even if the specific educational practices occurring during retention can be identified, a major weakness of retention research is ascertaining the causal effect of making students repeat a grade. As previously mentioned, random assignment of students to different treatment conditions is considered the most powerful research design to assess causality. The implausibility of randomly making academically challenged students repeat a grade leaves researchers only quasi-experimental designs to control for differences between retained and promoted students. The two types of general procedures appearing in the retention literature are matching and statistical adjustments. Although findings are considered to be superior if retained and promoted students are well matched, or if important variables known to be related to academic success can be incorporated into a regression equation, both of these procedures may still result in biased outcomes. One shortcoming pertains to the issue of regression artifacts. The second is referred to as the problem of "omitted" or "unmeasured" variables associated with selection biases.

All educators acknowledge that students of similar ability may not obtain identical scores on an examination. Students have good days and bad days when taking tests. Students with very low scores one year will likely have higher scores on the same test the following year. Similarly, children with extremely high scores one year will probably have somewhat lower scores when the exam is taken again. This general phenomenon is commonly referred to as “regression to the mean.” Campbell and Kenny (1999) caution that the process of matching or the use analysis of covariance methods may lead to findings that are simply regression artifacts. Studies of teacher initiated retention are problematic because teachers are more likely to make the lower performing students repeat a grade. A low performing retained pupil who is matched with a promoted student with a similar score may appear to be similar, but they still could have different levels of knowledge and ability. The retained pupil’s poor test results could be due to an abnormally low test score resulting from transient idiosyncratic factors during the day of the examination. The test results on which the teacher bases the retention decision may be due to a student having a low score which underestimates the child’s true ability. Students retained under such circumstances will likely obtain appreciably higher scores on next year’s exam because the initial test score was below the true level of student learning. Had the students required to repeat a grade been promoted, their test scores would still have been higher due to the regression to the mean phenomenon. Shepard (2002) and her colleagues (Shepard et al., 1996) argue that findings showing retention is associated with higher test scores result from regression artifacts. Critics of grade retention contend that the better test results observed among the retained students after being required to repeat a grade are mainly attributable to the regression to the mean, rather than better comprehension of the material. Statistical adjustments, such as entering the test score prior to the year of retention into a linear prediction model, may not adequately control for the regression effect. The lowest performing students, who are most likely to be retained, will probably report higher test scores at the end of the retention year.

Several methods have been suggested to determine the degree to which regression artifacts may be present in data. If test scores are available from several exams prior to retention, one can examine the test score means plotted over time to ascertain if student performance has been consistently declining. The presence of a steep decline in test results prior to the year of retention and a sharp increase following the grade repeated would imply a regression to the mean effect. If data prior to the year of retention exist, Campbell and Kenny (1999, pp.158-163) suggest performing a “time-reverse analysis” in which the values of the dependent variable and its covariate be reversed. That is, the value of the outcome measure after the retention year (T_2) would become the independent variable in a regression analysis while the initial value of the covariate (the value in T_1) becomes the dependent variable. Should the sign of the binary treatment variable reverse itself but remain similar in magnitude to that from the initial regression equation, a regression artifact is highly unlikely. The original estimated effect of the treatment is probably unbiased. Although critics of grade retention often argue that the positive effects of retention are due to regression artifacts, Campbell and Kenny (1999, pp.74-75) suggest that statistical adjustments likely underestimate the effect of treatments, especially when the pretest mean of the group given the treatment is smaller than that of the pretest value for the control group. Insofar as, prior to being retained, the mean test scores of students required to repeat a grade are almost always lower than those of the promoted students, analysis of covariance adjustments probably yield estimates lower than the true impact of the retention year on student academic performance.

An exacting methodologist could argue the major flaw of all retention research is that none have adequately addressed the issue of unmeasured or unspecified variables which affect the decision to retain a student. Some pupils are retained for academic reasons while others are held back because of a lack of emotional or behavioural maturity, which teachers or principals assume will retard future learning in the next grade. Those factors which lead to retention may also affect

academic learning outcomes. These unspecified causes will often result in a teacher recommending one student repeat the grade while another student with a similar test score will be allowed to progress to the next grade. If these unmeasured variables affect both the retention decision and performance on tests in later grades, the calculated net effect of grade retention will be biased, often referred to as the problem of “omitted variables.” Although matching or the use of test scores prior to retention as covariates will likely yield more accurate estimates of the impact of retention on academic achievement than not attempting to control for initial differences between retained and promoted children, neither procedure will result in unbiased findings.¹⁶

Over the last twenty years statisticians and econometricians have developed statistical procedures which can be used to help better address the problem of unmeasured variables. These techniques should allow educational researchers to obtain more accurate estimates of how making students repeat a grade, or any other educational intervention, influences their academic achievement. The current conceptual framework recommended to estimate causal effects is referred to as the “counterfactual account of causality.” Although this approach is highly technical in nature, Winship and Morgan (1999) provide a general overview of the basic issues and various analytical strategies which can be used to obtain more accurate assessments of the impact of educational practices on student outcomes. Only the most basic features of this approach is presented here. The counterfactual approach tries to estimate the effect of being placed in one group as opposed to another. In natural experiments individuals are nonrandomly assigned to a treatment group or a control group. For example, students would have an observable outcome measure if retained or promoted. The counterfactual approach attempts to answer the question of what would happen to the children who were retained if they were instead promoted to the next grade. One could also ask what the consequences of repeating a grade would be for students who were promoted to the next grade. Although students have potential outcome in either state, an outcome can be measured in only one state. For example, a retained student has an observed score when retained and an unobserved counterfactual outcome if placed in the control or promoted group. In the current context, those factors which lead to a student repeating a grade or being promoted will likely be associated with a student’s later academic achievement. Given that assignment to the treatment or control group will be correlated with the outcome variable of interest, standard ordinary least squares regression methods will not yield consistent estimates of the retention effect.

Numerous strategies have been suggested to reduce the correlation between the treatment and the outcome measure caused by assignment to the control or experimental group. An often used strategy is to create a “propensity score” or the probability that a person with certain characteristics will be assigned to the treatment group. One would use a large number of variables (Z_i) to compute the propensity score, $P(Z_i)$ of ending up in the treatment condition (here the retained group). The propensity score is like a mega-covariate for placement in the treatment group. Rosenbaum and Rubin (1985) then recommend matching the propensity score of a retained student with the closest propensity score of a promoted student in the control group. This approach assumes, however, that all variables which influenced placement into the treatment group are observed. If one assumes that unobserved or unspecified factors influence placement into the treatment or control group (a more reasonable assumption), Heckman (1978) suggests

¹⁶ Whereas the present paper concentrates exclusively on studies examining the impact of retention on academic achievement, another body of research focuses on the effect of grade retention and the completion of high school. The quality of research in these studies is generally higher than in the achievement literature. The consistency of findings from studies examining the impact of retention on school dropout behaviour suggests that making students repeat a grade will cause students to leave school without a diploma (Jimerson, Anderson, and Whipple, 2002). However, none of the studies examining the relationship between retention and dropping out of school have investigated the possibility that factors leading to retention also affect the likelihood of leaving school. Thus far, no one has investigated the possibility that models assessing the effect of grade retention on dropping out of school are misspecified because of unobserved variables affecting both retention status and school exit behaviour.

generating two variables, one for the likelihood of being placed in the treatment group and another for being selected into the control group. These two new selection effect variables can be entered into a regression equation along with the treatment variable and other control variables predicting the outcome measure of interest. Another strategy is to use instrumental variables which affect assignment into the experimental or control group, but does not directly influence the outcome measure. Although often used by economists, the instrumental variables approach suffers from certain shortcomings which may limit its usefulness (Winship and Morgan, 1999, pp.683-685). More rigorous analytical procedures to gauge the impact of retention on student performance require several measures of the outcome variable both prior to and following retention for students in the treatment and control groups are described by Winship and Morgan. The more advanced statistical modelling procedures suggested by the counterfactual approach to causality may yield new insights about the degree to which grade retention influences student learning outcomes.

Beginning with then President Clinton's (1998) appeal to end social promotion and culminating with the No Child Left Behind Act (2002) legislation, many state educational agencies and some local districts have mandated that students should be promoted only after mastering the basic skill requirements for their grade level. Political pressure to change school promotion policies will likely increase the number of children required to repeat a grade. Although critics have railed against the practice of retention, the present political climate in the United States will likely not change sufficiently in the near future to allow educational practitioners to modify recently implemented stringent promotion policies. However, the present situation may provide researchers an opportunity to more thoroughly evaluate retention practices. Whereas many previous studies were limited to a small number of students, the greater availability of retained students from more diverse racial and economic backgrounds should enable a better description of the consequences of retention. Given that not all districts will utilise the same remediation practices, the great variation in specific retention policies will allow researchers a better opportunity to identify which programs and strategies are more likely to enhance the academic standing of retained children. While such a suggestion may seem self-serving, educational researchers should take advantage of those situations which can help increase our understanding of retention practices and their implications for students. Even though many readers already have strong opinions on the retention issue, those individuals who acknowledge the limitations of the existing research should search for research settings which will enable us to further our comprehension of retention processes and their outcomes.

REFERENCES

Note: References marked with a single asterisk (*) indicate studies appearing in the Holmes (1989) meta-analysis.
References with two asterisks (**) are cited in the Jimerson (2001) meta-analysis.

- *Abidin, R. P., Golladay, W. M., and Howerton, A. L. (1971). Elementary school retention: An unjustifiable, discriminatory, and noxious educational policy. *Journal of School Psychology*, 9(4), 410-417.
- **Alexander, K. L., Entwisle, D. R., and Dauber, S. L. (1994). *On the Success of Failure: A Reassessment of the Effects of Retention in the Primary Grades*. New York: Cambridge University Press.
- Alexander, K. L., Entwisle, D. R., and Dauber, S. L. (2003). *On the Success of Failure: A Reassessment of the Effects of Retention in the Primary Grades* (2nd ed.). New York: Cambridge University Press.
- Anfinson, R. D. (1941). School progress and pupil adjustments. *The Elementary School Journal*, 41(7), 507-514.

- Archer, G. A. (1967). A Study of Nonpromotion relative to normal grade expectancy in selected Catholic elementary schools in Illinois. Unpublished doctoral dissertation, Loyola University, Chicago. Cited in C. T. Holmes (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard and M. L. Smith (eds). *Flunking grades: Research and policies on retention*. London: The Falmer Press.
- Campbell, D. T. and Kenny, D. A. (1999). *A Primer on Regression Artifacts*. New York: Guilford Press.
- Campbell, D. T. and Stanley, J. C. (1966). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- *Chansky, N. M. (1964). Progress of promoted and repeating grade 1 failures. *Journal of Experimental Education*, 32(3), 225-237.
- Clinton, W. (1998). State of the Union Address (January 1, 1999). Washington, DC: U.S. Government Printing Office.
- Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences* (Revised ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Coleman, J. S. and Karweit, N. L. (1972). *Information Systems and Performance Measures in Schools*. Englewood Cliffs, NJ: Educational Technology Publications. [The Rand Corporation]
- Cosden, M., Zimmer, J., Reyes, C., and del Rosario Gutierrez, M. (1995). Kindergarten practices and first-grade achievement for Latino Spanish-speaking, Latino English-speaking, and Anglo students. *Journal of School Psychology*, 33(2), 123-141.
- Darling-Hammond, L. and Falk, B. (1997). Using standards and assessments to support student learning. *Phi Delta Kappan*, 79(3), 190-199.
- Dawson, P. (1998). A primer on student grade retention: What the research says. *NASP Communique*, 26(8), 28-30. [Online] <http://www.nasponline.org/publications/cq268retain.html> [June 18, 2004].
- **Dennebaum, J. M. and Kulberg, J. M. (1994). Kindergarten retention and transition classrooms: Their relationship to achievement. *Psychology in the Schools*, 31(1), 5-12.
- *Dobbs, V. and Neville, D. (1967). The effect of nonpromotion on the achievement of groups matched from retained first graders and promoted second graders. *Journal of Educational Research*, 60(10), 470-475.
- Dunkin, M. J. (1996). Types of errors in synthesizing research in education. *Review of Educational Research*, 66(2), 87-97.
- **Ferguson, P. C. (1991). Longitudinal outcome differences among promoted and transitional at-risk kindergarten students. *Psychology in the Schools*, 28(2), 139-146.
- **Ferguson, P. C. and Mueller-Streib, M. (1996). Longitudinal outcome effects of non-at-risk and at-risk transitions first-grade samples: A follow-up study and further analysis. *Psychology in the Schools*, 33(1), 38-45.
- **Hagborg, W. J., Masella, G., Palladino, P. and Shepardson, J. (1991). A follow-up study of high school students with a history of grade retention. *Psychology in the Schools*, 28(4), 310-317.
- Harrington-Lueker, D. (1998). Retention vs. social promotion. *The School Administrator*, 55(7), 6-12.

- Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46(4), 931-61.
- Hess, A. G. (1995). *Restructuring Urban Schools*. New York: Teachers College Press.
- Heubert, J. P. and Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington D.C.: National Academy Press.
- Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard and M. L. Smith (eds). *Flunking grades: Research and Policies on Retention*. London: The Falmer Press.
- Holmes, C. T. and Matthews, K. M. (1984). The effects of non-promotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research*, 54(2), 225-236.
- Jackson, G. B. (1975). Research evidence on the effects of grade retention. *Review of Educational Research*, 45(4), 613-635.
- Jacobs, B. A. and Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *The Review of Economics and Statistics*, 86(1), 226-244.
- Jimerson, S. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30(3), 420-437.
- **Jimerson, S. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37(3), 243-272.
- Jimerson, S. (2004). Is grade retention educational malpractice? Empirical evidence from meta-analyses examining the efficacy of grade retention. In H. Walberg, A. J. Reynolds and M. C. Wang (eds). *Can Unlike Students Learn Together?* Greenwich, CT: Information Age Publishing, Inc.
- **Jimerson, S., Carlson, E., Rotert, M., Egeland, B., and Sroufe, L. A. (1997). A prospective, longitudinal study of correlates and consequences of early grade retention. *Journal of School Psychology*, 35(1), 3-25.
- Jimerson, S., Anderson, G. E., and Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relationship between grade retention and dropping out of high school. *Psychology in the Schools*, 39(4), 441-457.
- **Johnson, E. R., Merrell, K. W., and Stover, L. (1990). The effects of early grade retention on the academic achievement of fourth-grade students. *Psychology in the Schools*, 27(4), 333-338.
- *Kamii, C. K. and Weikart, D. P. (1963). Marks, achievement, and intelligence of seventh graders who were retained (nonpromoted) once in elementary school. *Journal of Educational Research*, 56(9), 452-459.
- Kaplan, R. M. and Saccuzzo, D. P. (1997). *Psychological Testing: Principles, applications, and issues* (4th ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.
- Kaestle, C. F. (1993). The awful reputation of educational research. *Educational Researcher*, 22(1), 25-31.
- Karweit, N. L. (1992). Retention policy. In M. C. Alkin (ed) *Encyclopedia of Educational Research* (6th ed.), New York: Macmillan.

- Karweit, N. L. (1999). *Grade retention: Prevalence, timing, and effects* (Report No. 33). Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.
- *Leinhardt, G. (1980). Transition rooms: Promoting maturation or reading education? *Journal of Educational Psychology*, 72(1), 55-61.
- Lorence, J., Dworkin, A. G., Toenjes, L. A., and Hill, A. N. (2002). Grade retention and social promotion in Texas, 1994-1999: Academic achievement among elementary school students. In D. Ravitch (ed) *Brookings Papers on Education Policy 2002*. Washington DC: The Brookings Institution.
- Mabry, L. (1995). Review of the Metropolitan Readiness Tests, 5th ed. In J. C. Impara and J. C. Conoley (eds). *Twelfth Mental Measurements Yearbook*. Lincoln, NE: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln.
- **Mantzicopoulos, P. Y. (1997). Do certain groups of children profit from early retention? A follow-up study of kindergartners with attention problems. *Psychology in the Schools*, 34(2), 115-127.
- **Mantzicopoulos, P. and Morrison, D. (1992). Kindergarten retention: academic and behavioral outcomes through the end of second grade. *American Educational Research Journal*, 29(1), 182-198.
- *May, D. C. and Welch, E. L. (1984). The effects of developmental placement and early retention on children's later scores on standardised tests. *Psychology in the Schools*, 21(3), 381-385.
- Maxwell, S. E. and Delaney, H. D. (1990). *Designing experiments and analyzing data*. Belmont, CA: Wadsworth Publishing.
- Maxwell, S. E., O'Callaghan, M. F., and Delaney, H. D. (1993). Analysis of covariance. In L. K. Edwards (ed) *Applied Analysis of Variance in Behavioral Science*, New York: Marcel Dekker.
- **McCombs Thomas, A., Armistead, L., Kempton, T., Lynch, S., Forehand, R., Nousiainen, S., Neighbors, B., and Tannenbaum, L. (1992). Early retention: Are there long-term beneficial effects? *Psychology in the Schools*, 29(4), 342-347.
- **McCoy, A. R. and Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology*, 37(3), 273-298.
- **Meisels, S. J. and Liaw, F. (1993). Failure in grade: Do retained children catch up? *Journal of Educational Research*, 87(2), 69-77.
- National Association of School Psychologists [NASP]. (2003). Position statement on student grade retention and social promotion (revised April 12, 2003). Retrieved June 18, 2004, from http://www.nasponline.org/information/pospaper_graderetent.html.
- *Niklason, L. B. (1984). Nonpromotion: A pseudoscientific solution. *Psychology in the Schools*, 21(4), 485-499.
- No Child Left Behind Act of 2001, 20 U.S.C. § 6301 (2002).
- Owings, W. A. and Magliaro, S. (1998). Grade retention: A history of failure. *Educational Leadership*, 56(1), 86-88.
- *Peterson, S. E., DeGracie, J. S., and Ayabe, C. R. (1987). A longitudinal study of the effects of retention/promotion on academic achievement. *American Educational Research Journal*, 24(1), 107-118.

- **Phelps, L., Dowdell, N., Rizzo, F. G., Ehrlich, P., and Wilczenski, F. (1992). Five to ten years after placement: The long-term efficacy of retention and pre-grade transition. *Journal of Psychoeducational Assessment*, 10(3), 116-123.
- **Pierson, L. and Connell, J. P. (1992). Effect of grade retention on self-system processes, social engagement, and academic performance. *Journal of Educational Psychology*, 84(3), 300-307.
- Pomplun, M. (1988). Retention, the earlier the better? *Journal of Educational Research*, 81(5), 281-287.
- Potter, L. (1996). Examining the negative effects of retention in our schools. *Education*, 117(2), 268-270, 250.
- **Reynolds, A. J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis*, 14(2), 101-121.
- **Reynolds, A. J. and Bezruczko, N. (1993). School adjustment of children at risk through fourth grade. *Merrill-Palmer Quarterly*, 39(4), 457-480.
- Reynolds, A. J., Temple, J., and McCoy, A. (1997). Grade retention doesn't work. *Education Week*, 17(3) [September 17], 36.
- Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 39(1), 33-38.
- **Rust, J. O., and Wallace, K. A. (1993). Effects of grade level retention for four years. *Journal of Instructional Psychology*, 20(2), 162-166.
- *Sandoval, J. and Fitzgerald, P. (1985). A high school follow-up of children who were nonpromoted or attended a junior first grade. *Psychology in the Schools*, 22(2), 164-170.
- Schuyler, N. B. and Matter, M. K. (1983). To retain or not to retain: Should achievement be your guide? Paper presented at the annual meeting of the American Educational Research Association, Montreal. Cited in C. T. Holmes (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard and M. L. Smith (eds) *Flunking Grades: Research and Policies on Retention*. London: The Falmer Press.
- Seltzer, M. H., Frank, K. A., and Bryk, A. S. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation and Policy Analysis*, 16(1), 41-49.
- Shepard, L. (2000). Cited in Ending social promotion by Debra Viadero. *Education Week on the Web*. (March 15, 2000). Retrieved on June 1, 2004, from [Online] http://www.edweek.org/ew/ewstory.cfm?slug=27.social.h19andkeywords=Ending_per_cent20Social_per_cent20Promotion.
- Shepard, L. (2002). Comment on grade retention and social promotion in Texas, 1994-1999: Academic achievement among elementary school students. In D. Ravitch (ed) *Brookings Papers on Education Policy 2002*. Washington DC: The Brookings Institution.
- Shepard, L. (2004). Understanding research on the consequences of retention. In H. Walberg, A. J. Reynolds and M. C. Wang (eds) *Can Unlike Students Learn Together?* Greenwich, CT: Information Age Publishing, Inc.
- *Shepard, L. and Smith, M. L. (1987). Effects of kindergarten retention at the end of first grade. *Psychology in the Schools*, 24(4), 346-357.

- Shepard, L. A. and Smith, M. L. (1990). Synthesis of research on grade retention. *Educational Leadership*, 47(8), 84-88.
- Shepard, L. A., Smith, M. L., and Marion, S. F. (1996). Failed evidence on grade retention. [Review of the book *On the success of failure: A reassessment of the effects of retention in the primary grades.*] *Psychology in the Schools*, 33(3), 251-261.
- Southard, N. A. and May, D. C. (1996). The effects of pre-first grade programs on student reading and mathematics achievement. *Psychology in the Schools*, 33(2), 132-142.
- Stoner, G. (1995). Review of the Metropolitan Readiness Tests, 5th ed. In J. C. Impara and J. C. Conoley (eds), *Twelfth Mental Measurements Yearbook*. Lincoln, NE: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln.
- Tanner, C. K. and Galis, S. A. (1997). Student retention: Why is there a gap between the majority of research findings and school practice? *Psychology in the Schools*, 34(2), 107-114.
- Thorndike, R., M. (1997). *Measurement and Evaluation in Psychology and Education* (6th ed.). Upper Saddle River, NJ: Prentice Hall
- U.S. Department of Education. (1999). *Taking Responsibility for Ending Social Promotion*. Washington, DC: U.S. Government Printing Office.
- Vander Weele, M. (1994). *Reclaiming our Schools: The Struggle for Chicago School Reform*. Chicago: Loyola University Press.
- Wilson, M. (1990). Review of the book, *Flunking Grades: Research and Policies on Retention*. *Educational Evaluation and Policy Analysis*, 12(2), 228-230.
- Winship, C. and Morgan, S. L. (1999). The estimation of causal effects from observational data. In K.S. Cook and John Hagan (eds) *Annual Review of Sociology Vol. 25*. Palo Alto, CA: Annual Reviews.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative Methods for Research Synthesis*. Beverly Hills, CA: Sage.
- Wright, J. B. (1979). The measured academic achievement of two groups of first grade students matched along five variables when one group has been retained. Unpublished doctoral dissertation, Loyola University, Chicago. Cited in C. T Holmes (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard and M. L. Smith (eds) *Flunking Grades: Research and Policies on Retention*. London: The Falmer Press.

Successful pedagogies for an Australian multicultural classroom

Carlene Winch-Dummett

Flinders University, School of Education aspwinch@netspeed.com.au

A study undertaken in two primary schools and a pre-school in a multicultural urban area of NSW identified five pedagogical topics that incorporated successful teaching practices and processes that accord with the Delors Report recommendations. The report recommended that education for the future should be organised around the four pillars of learning, namely, learning to be, learning to do, learning to know and learning to live together. The five identified topics were: lesson organisation, lesson outcomes, teacher communication, teaching strategies, and cultural inclusion.

In order to address the needs of students, the five topics have been grouped into two student-focused themes: (a) the learning environment and (b) teacher-student communication. These themes are discussed from a theoretical perspective and illustrated with examples derived from research.

Cultural inclusion, pedagogies, *Delors Report*, learning environment, teacher-student communication, lesson organisation, multicultural classroom

INTRODUCTION TO THE STUDY

A study undertaken in 2002 in two New South Wales Department of Education and Training (NSW DET) primary schools, School A and School B, and a pre-school approved by the New South Wales Department of Community Services (NSW DOCS), School C, aimed to identify the pedagogical processes that target the recognition, importance and sharing of cultural diversity in teaching and learning as strategies for learning to live together in an Australian multicultural classroom. These pedagogical processes refer to teaching processes and practices that respond to the Four Pillars of Learning described as learning to be, learning to do, learning to know and learning to live together, in the UNESCO Delors Report *Learning: the Treasure Within* (Delors, 1996).

The schools were in an urban area of NSW which was characterised by high immigration from a diversity of countries and cultures, particularly since the end of World War II. A teacher from each of the primary schools was chosen for observation because each had been designated as a teacher of excellence by her peers and senior educational officials, and one had received an award from the NSW Department of Education and Training (DET) for teaching excellence. The teachers have been identified as Teacher A, who taught English as a Second Language (ESL) Kindergarten class in School A; and Teacher B, who taught a combined Years 1 and 2 mainstream class, in which six of the 26 students were from non-English speaking backgrounds (NESB), two mornings each week. Teacher B's duties also included taking one of her Year 2 students and two students from other Year 1 and Year 2 classrooms for one-on-one withdrawal lessons, or in-class assistance in ESL.

The study undertaken in School C, the pre-school, was limited to the classroom environment which was recognised by the local community and the NSW DOCS as meeting the needs of the local multicultural community. It did not include observations of the teaching processes and

practices of the pre-school teacher. However, the pre-school was observed to have a broad range of cultural resources, many of which were borrowed from the local Children's Resource Centre that was funded by the Commonwealth Government Supplementary Services (SUPS) program that targeted children with special needs.

The research procedure used to identify the successful teaching processes and practices involved a synchronic analysis of the lessons of Teacher A and Teacher B, and a categorical analysis of the teaching practices of Teacher A and Teacher B. A categorical analysis of the resources of the classroom environment of the pre-school was made in order to identify the ways in which the classroom environment might enhance teaching processes and practices in multicultural classrooms.

FINDINGS FROM THE STUDY

The findings from the analyses into the successful teaching processes and practices of Teacher A and Teacher B and the learning environment of School C were combined to form five pedagogical topics. These five topics drew together specific teaching methods that addressed the lesson and syllabus outcomes of the relevant educational levels, and demonstrated their appropriateness for inclusion in classrooms where the amount of cultural diversity was a variable factor. The five identified pedagogical topics that incorporated successful teaching process and practices were described as follows:

- (a) lesson organisation which included time and place;
- (b) lesson outcomes which were informed by syllabus guidelines and included teaching processes such as scaffolding and transfer;
- (c) teacher communication which included verbal communication and non-verbal communication such as gestures, proxemics and paralanguage;
- (d) teaching strategies which included questioning, feedback, holistic learning and teaching to multiple intelligences; and
- (e) cultural inclusion which acknowledged culturally specific activities and beliefs, promoted values, and established an ethos of respect for cultural diversity.

Since the research into the successful teaching processes and practices was focused on the teachers' performances, it was appropriate that the discussion of the implementation of these teaching processes and practices should be learner-focused. The reason for this was that even the youngest learners were culturally nurtured individuals with their own pre-learned knowledge, and since teachers were also the products of their cultural backgrounds, it was not possible for teachers to conclude, with certainty, that learners would always reflect, understand or attain the meanings that they were attempting to share or provide.

A learner focused discussion of the application of the components of the five successful pedagogical topics required a reorganisation of the components from those that reported on the teachers' performance to those that reflected the needs of learners. Therefore, two major student-focused themes were devised that incorporated the components of the five pedagogical topics. These two themes were:

- (a) learning environment.
- (b) teacher-student communication.

The theme 'learning environment' is comprised of: (a) scaffolding and transfer as instruction techniques; (b) the theory and application of holistic learning; and (c) teaching strategies that specifically target the individual abilities and talents of children, and that are consistent with Gardner's (1999) theory of multiple intelligences. Teacher-student communication is also sub-

divided into verbal and non-verbal communication. Verbal communication consists of the effect of prior language learning of children at the entry level into formal education, language as a symbolic system, inference, questioning and feedback. Non-verbal language includes features such as gesture, proxemics, voice intonation and gaze. An overview of the applications and implications of teaching strategies, organised under the themes of learning environment and teacher-student communication, consists of discussions of relevant theoretical literature that are illustrated with examples from the research conducted in the primary schools and the pre-school.

The Learning Environment

The research undertaken in School A and School B exemplifies a range of attitudes to early education that varies from the attitudes of parents who encourage children to explore learning opportunities such as libraries, clubs and other extra curricular and after school activities, to those of parents who have minimal interaction with their children because of work or domestic commitments. This includes parents who insist that children at the entry level into early formal education are too young to be concerned with serious learning. Similarly, parental participation in school life varies. Some parents are frequent visitors to the classroom and participate in activities such as developmental play and excursions. Other parents are interested in their children's progress, but do not actively participate in school activities, except to attend parent-teacher meetings that are associated with their children's progress. A minority of parents believe that early formal education can accommodate the child's interests rather than forcing the child to acquire new concepts or skills, especially if the child has learning problems in the development of new concepts or skills.

The challenge for the teacher is how to accommodate pre-existing or culturally modified learning methods and the parental expectations and attitudes that attend the young learner in early schooling. This research study identifies several pedagogical processes that address diverse learning experiences, but it also reveals that teachers need to work in supportive school programs and organisational structures if these approaches are to be successful.

The pedagogical processes that are identified are those that are conducive to the learning environment and address the extant cultural backgrounds of the students. The NSW Department of Education has recommended the following educational outcomes:

- (a) scaffolding and transfer of knowledge.
- (b) holistic learning.
- (c) teaching to multiple intelligences.

Scaffolding and the transfer of knowledge are methods of instruction; holistic learning is the organisation and integration of the knowledge and skills to be acquired; and teaching strategies that address multiple intelligences are programs and processes that specifically target abilities, talents and pre-existing skills of children that enable them to acquire knowledge and skills.

Scaffolding and Transfer of Knowledge

Scaffolding is the structured building of competencies on knowledge already possessed by the learner. It involves the connection between familiar and new information. Bransford, Brown and Cocking (1999) claim that tasks associated with scaffolding should motivate the learner, simplify the task, particularly the number of steps involved, focus on the goal or outcome, take note of significant discrepancies between the desired goal and the learner's achievements, control frustration and provide a model of the skill to be performed. Transfer of knowledge is similar to scaffolding, but is associated with using acquired knowledge in new situations and transferring knowledge from one context to another. Successful transfer requires mastery of the original

subject. Consequently knowledge learned by rote or memorised is not as effectively transferred as that which is understood.

Scaffolding and the transfer of knowledge are teaching strategies that require planning that is focused on the specific lesson outcomes and the long term educational program. Scaffolding may be problematic in classrooms where some children are frequently absent from school and therefore may miss stages in the learning process. The teacher can redress this difficult situation by including exercises that involve transfer as well as scaffolding.

Teacher B, who taught a combined Year 1-2 class, also provided support to ESL students, and relied on transfer rather than scaffolding to develop skills. She was employed part time and this precluded the use of an incremental scaffolding method of teaching. Since she taught literacy, numeracy, arts and crafts in the integrated Year 1-2 class, she collaborated with the class teacher to develop thematic approaches to the subject matter. This strategy made it possible for knowledge to be transferred from one learning discipline to another and from one teacher to the other. For example, in Term 3 of 2002, science, literacy and the arts and crafts lessons were organised around the theme of frogs. The thematic approach facilitated the transfer of skills and knowledge across the curriculum. In the individual support lessons for ESL students, she addressed specific skills being taught in the child's home class within the content of a different subject. This generated a transfer of learning from one context to another. Bransford et al. (1999) (1999) maintained that knowledge transfer was more effective when the subject was taught in more than one context.

Bransford et al. (1999) recognised the influence of culture on learning. They suggested that school failure might be the result of a mismatch between what was learned at home and what was learned at school because different meanings attached to cultural knowledge could affect transfer.

Holistic Learning

Holistic learning refers to the organisation of knowledge and skills to be addressed in school education. Holistic learning is an approach to learning that is all inclusive in terms of subject areas and the allocation of time to the acquisition of skills and knowledge. This type of learning presupposes an integrated curriculum and flexibility in lesson structure and the time allocated for learning. It breaks down artificial barriers between subject areas and seeks to explore subjects in breadth and depth. Holistic learning presupposes that lesson structures reflect the fluid nature of learning and are not interrupted by arbitrary time constraints. Holistic learning encompasses not only subjects that are measurable, but also a more spiritual or ethical depth to learning, that is, a recognition of the values that underpin human relations in society. These values extend beyond person to person interactions to encompass the care of the environment in all its manifestations such as animal life, nature and heritage.

Holistic learning may contradict Western notions of compartmentalisation of life and work. It confronts a world view where time is a linear concept focused on progress and therefore a future that is not within an individual's grasp or understanding. It advocates an interconnectedness of experience in which values are intrinsic to human activity and in which reflection on past experience is a positive means of informing the present and predicting the future. It permits children to explore the breadth of their subjects and to make connections between contemporary and past learning.

According to Duffy (1994), holistic learning includes a world view and a focus on the humaneness of the individual. Duffy believes that a world view should reflect the connectedness, wholeness and meaning of experience; the humaneness of the individual should avoid dichotomies between mind and body, intellect and emotion, and rationality and intuition. He claims that an integrated curriculum should draw together these issues, and include a spiritual

dimension to learning. The form of spirituality that Duffy envisages is not one that is based on religious belief but one that reflects on the individual in all of his or her being, that is, with cosmology, the nature of being, and a connection to the environment including ethics that contribute to humaneness. This point is supported by Teasdale (1999, p.82) who suggested that:

The whole area of cultural studies, civics, moral education, learning to live together, peace education, and spirituality cannot be taught by putting it into a box and trying to slot it into an already overcrowded curriculum. It is necessary to find ways of infusing these elements across all aspects of the curriculum.

Hall (1976) examined the Western world linear view of time, one that could be controlled and manipulated and that was focused on an immediate future. Hall contrasted this Western organisation of time with the Arab view that was polychronic. The polychronic view of time was illustrated by the fact that Arabs were highly sociable, rarely alone, and tended to carry on conversations with more than one individual simultaneously. Hall made the point that in cultures where time and action were not compartmentalised, the participants had a broader understanding of context than Westerners whose context for action was limited to one stage of a process at a time, that is, a linear view of time.

Adams (1968) described the American concept of time, (and this could be extrapolated to include all Western societies) as future focused. He made the point that time was viewed as cyclical in Korea, with the future a kind of repetition of the past. Adams believed that the different attitudes of the Americans and the Koreans to time were reflected in their attitudes and cultures. Westerners were preoccupied with dominating culture, whereas, Koreans lived in harmony with culture and therefore would prefer to conform rather than to create. This concept of time as being non-linear, reflecting the past, the present and the future, was discussed by Teasdale and Teasdale (1999) who described the way in which Aboriginal knowledge was reinforced by a focus that transcended time. Aboriginal people conceived time as past continuous. It was cyclical rather than linear whereby the past was part of the present and the future.

Holistic education has the capacity to develop the quality that Duffy (1994) describes as humaneness. This may be achieved through values education. In their overview of the current debate on values in the curriculum of Australian schools, Aspin, Chapman and Klenowski (2001) discuss the relevance of values education and its embodiment in subjects throughout the curriculum. They acknowledge that educators are increasingly receptive to the argument that values are not appropriately taught as a separate discipline. Consequently, they examine the ways in which values relate to the understanding and appreciation of their own cultures and those of others, as well as their value to the community.

Teaching to Multiple Intelligences

The teaching strategies are required to address the multiple intelligences of students and acknowledge the specific talents, abilities and aptitudes of individual students in order to achieve optimal accessibility to knowledge acquisition, understanding, , and skills development. Gardner (1999) reviewed his theory of multiple intelligences and described eight autonomous aptitudes or intelligences through which individuals learned. Originally seven intelligences were described (Gardner, 1983). These were: linguistic, logical, spatial, bodily kinaesthetic, musical, interpersonal, and intrapersonal intelligences. More recently Gardner included an eighth intelligence, a naturalistic intelligence; and also canvassed the possibility of further intelligences such as spiritual and existential intelligences (Gardner, 1999). Spiritual intelligence referred to mythology, religious beliefs and their expression in the arts. Existential intelligence referred to issues such as the place of the individual in the cosmos. Gardner acknowledged that such spiritual and existential intelligences did not meet his criteria for description as intelligences that were informed by logical analysis, developmental psychology and psychological research.

Gardner (1983) claimed that an intellectual competence must include problem solving skills that enable individuals to solve problems and to find or create products. A prerequisite for a theory of multiple intelligences was that the intelligences were required to reflect the whole range of abilities valued by human cultures. Gardner (1983) provided eight signs that represented the criteria for a faculty or an aptitude to be considered intelligence. These eight signs are:

- (a) potential isolation by brain damage,
- (b) the existence of idiot savants, prodigies, and other exceptional individuals,
- (c) an identifiable core of operation or set of operations,
- (d) a distinctive developmental history, along with a definable set of expert 'end-state' performances,
- (e) an evolutionary history and evolutionary plausibility,
- (f) support from psychological tasks,
- (g) support from psychometric findings, and
- (h) susceptibility to encoding in a symbol system.

Critics of Gardner's theory of multiple intelligences such as Alix (2000) argued that it did not meet criteria for theoretical adequacy, or methodological validity and reliability in application. Nevertheless, as a strategy for accessing the inherent talents and abilities of children to generate learning, Gardner's theory of multiple intelligences changed the direction of instruction from one that was focused predominantly on linguistic and mathematical conceptualisation, to one in which the application of these skills and others were integrated within a broader learning context. The theory of multiple intelligences supported the view that there were other valid and significant areas of learning, performing and ultimately working in the vocational sense, such as the arts or team membership. The contemporary dynamic social changes described by Delors (1996), and the consequent vocational demands for a versatile workforce covering a broad range of professions, demanded a wide range of skills and knowledge as well as linguistic, mathematical and scientific competencies. The implementation of Gardner's theory of multiple intelligences should be fostered for both the immediate learning process of students, and as a preparation for a future vocational application.

The following brief description of Gardner's (1999) eight intelligences includes examples of relevant teaching strategies used by Teacher A in the ESL Kindergarten class in School A, Teacher B in the integrated Year 1-2 class in School B; and the learning environment of School C, that according with strategies for teaching to multiple intelligences described by Armstrong (2000).

Linguistic intelligence is the capacity to learn and appreciate language, and to use language as a means of achieving designated goals. Examples from the schools studied are: storytelling, brainstorming and video-making; recall writing; and the publication of class work on the classroom computer. These examples are considered open-ended and therefore accessible to a wide range of learners.

Logical-mathematical intelligence involves the capacity to analyse, investigate and carry out mathematical operations. Strategies include using calculations, classifications and categorisations such as in numeracy exercises relevant to the literacy lesson; and counting backwards and forwards from the present day's date.

Spatial intelligence involves the capacity to recognise and manipulate space, in both its widest parameters such as that used by aeronautical engineers, and the more confined spaces of artists and artisans. Teaching strategies include reflection in spoken or written narrative; creating a story

or craft exercise to be based on imagination or abstract reasoning; and a reading program that includes graphics and colour cues.

Kinaesthetic intelligence concerns the use of the body to produce items such as arts and crafts, sports, and aesthetic activities, for example, (a) walking and skipping along a snake painted on the floor, (b) hands-on learning where children use construction pieces such as blocks, connecting blocks, beads, paper and scissors, and (c) other opportunities for full or part body involvement in action songs and numeracy exercises.

Musical intelligence is the capacity to appreciate, compose and perform musical patterns. Examples include the opportunity to examine musical instruments from a variety of cultural backgrounds, and a reading program which reinforces cultural differences through songs and chants.

Interpersonal intelligence involves the ability to perceive and act upon the intentions and motivations of others. Teaching strategies for interpersonal intelligence are increasingly important in an environment where team work draws experts and artisans from a variety of disciplines. It is also important for developing cross-cultural understanding in the classroom and in the wider society. Examples include peer sharing and group work, and opportunities for interaction with other children through organised activities or outdoor play.

Intrapersonal intelligence is the capacity to understand oneself, such as one's fears, motivations and aspirations, and the application of this understanding in regulating one's life. Examples include one minute reflection periods, recalling and relating experiences, discussing the new task or previous work, and free time to pursue topics of interest.

Naturalistic intelligence is the capacity to recognise and classify species of fauna and flora in one's environment. Strategies include excursions around the local neighbourhood, a visit to a nature park, a section of the classroom dedicated to natural history, and the encouragement of children to take care of their classroom and school environment.

Armstrong (2000) argues that, according to multiple intelligences theory, instructional objectives may be learned through the eight different intelligences and therefore it is possible for teachers to assess children's learning through methods reflecting any one of these eight intelligences. Armstrong presents a detailed model for assessment which he describes as eight ways of assessment. This model provides assessment tools consistent with a competence setting that is most relevant to a learner's individual intelligences. Armstrong recommends: (a) individual student portfolios with assessment based on student work samples; (b) opportunities for students to be assisted in developing their cognition through reflection on their work; (c) communication with all stakeholders in the student's education; (d) group work in the classroom; and (e) a set of criteria for comparison of the student's progress across the grades.

In Teacher A's classroom, assessment reflected Armstrong's recommendations. Portfolios were used by School A in 2002 in accordance with the NSW DET guidelines for outcomes-based assessment (NSW DET 2002). The portfolios were managed by the teachers of the home class kindergartens. They included samples of work carried out in the ESL lessons. In-school literacy and numeracy testing was undertaken in Year 2, and the NSW DET prescribed basic skills tests were carried out in Years 3 and 5. Parent-teacher meetings on student progress involved conferences with the home class teacher, the ESL teacher, parents and student as the contributing stakeholders in the student's education.

Teacher-Student Communication

Effective learning requires mutual understanding in any communication between the teacher and student(s). In classrooms where the teacher and children share the same first language there is

always the possibility that the spoken message may contain subtleties, inferences and ambiguities that affect its reception. In classes where learners come from a non-English speaking background (NESB) students have a prior-learned non-native language that they may use to reason, a bank of culturally learned attitudes and non-verbal signifiers, and limited English, it is an imperative that careful consideration is given to how information is imparted and received by the teacher, the parent or caregiver, and the student.

Verbal Communication

Verbal communication refers to any communication that uses words. This includes the spoken or written word in all its forms. The effective use of verbal communication in primary schools is essentially the spoken word. Language is a semiotic system, a system of signs. It is one of many semiotic systems that cultures use to inform members. Geertz (1973) claims that it is through the trafficking in signs that meaning can be transferred among individuals in any society.

Words, that are symbols used in language, are arbitrary signs until they acquire meaning. In the application of Saussurean linguistic theory, described by Hawkes (1977), words are signs that acquire meaning when used in relation to other words. When these words are strung into a syntax which, because of the selection of signs and their relation to one another produces meaning, they form sentences and utterances with extensive potential depending in scale on the richness of known words. While utterances are restricted by the time it takes for their presentation in a sentence, the written word may not be as severely constrained as evidenced in Asian scripts, or even contemporary Western poetry, or advertising where the relationships between signs may be accessed from a number of directions thereby enriching meaning through connotation (Hawkes, 1977).

Since children, and in fact persons at any age from a NESB background, commence formal education in a Western focused classroom in Australia with a prior learned system of communication, it is important for teachers to understand some of the issues associated with communication in order to generate a general approach to communication with NESB learners. In their discussion of the role of everyday discourse and social interaction in reinforcing values, ideologies and patterns of social organisation, Garret and Baquedano-Lopez (2002) refer to the studies by Schieffelin and Ochs (1986) that young children's socialisation activities with their care-givers leads to the acquisition of social and linguistic skills, and a culturally specific world view.

According to Brown (1974), children between the ages of 18 months and five years acquire a syntactic rule system that enables them to generate linguistic communication through the construction of numerous sentences comprehensible to a variety of community members. Brown argues that children do not learn to remember sentences; they gradually recognise and extract a set of rules of construction which provides them with the skills to produce further sentences. These sentence constructions or patterns are copied from the utterances of adults.

However, the adults do not necessarily understand the abstract constructions of the utterances. Therefore, it may be extrapolated from Brown's argument that since the rules of speech, and the patterns of speech by which children imitate the utterances of adult members of the immediate family, can be learned before commencing formal education, the teacher in a multicultural classroom may be confronted with a variety of prior-learned syntactic, tonal and rhythmic patterns of language. Moreover, these verbal qualities are likely to be accompanied by culturally acquired gestures and proxemics. Consequently, communication is the most immediate and significant of all teaching processes and practices from the point of the child are perhaps the most challenging for the teacher.

Questioning is perhaps the most difficult form of communication in which teachers engage with their students. In verbal communication where the teacher uses directives, assertions and explanations, the teacher has control over the communication that is chosen to convey the message, its context and its purpose within the assumed comprehension level of the learner. Questioning differs from information presenting utterances in that the communication is controlled more by the student's reception of the question, the student's interpretation of the question, and the student's attitude to the question, all of which are culturally learned. Rhetorical questioning such as "What are you doing?" which might or might not require an answer, is often associated with classroom management and maybe problematic for the NESB learner.

Heath's (1982) cross cultural studies of questioning by white middle class teachers in African-American classrooms, and also with their own children, confirm the complexity in the application of questioning strategies. Heath's findings revealed that children from infancy in white middle class societies learned Western type question rituals using identification type questions such as "What is that?". African-American questioning was used mainly in story-telling, and was more typically used for comparative purposes such as "What's that like?" Heath noted that African-American children did not always have experience in dealing with so-called 'how' and 'why' type questions. Other types of questions that were problematic to African-American children were those that were implied directives, such as "Why don't you get it from the back shelf?", and questions that required learned skills in narrative analysis, such as "Who will help Tim find his way home?" Hansen (1979) referred to Dickeman's (1971) observations that in some societies children learnt early in life not to divulge information, and in other social settings, children refused to answer a question if they believed that the questioner already knew the answer.

Hansen (1979) made the point that it was not possible to understand the exchanges between members of a classroom without knowledge of the cultural repertoire of the participants. Moreover, it was not always possible for a teacher to address the issues of communication with every member of a multicultural classroom. However, the teacher could be cognisant of cultural differences and could make adjustments to the questioning technique accordingly. Heath suggested a two-way path for questioning. This technique recommended that teachers should learn about the forms and functions of questions to which children were accustomed, and should encourage learners to acquire the skills to respond to questions according to the rules of classroom question usage.

Feedback referred to the communication by teachers of their evaluation of the learner's work to the student. Feedback should be used in classrooms as reinforcement of correct responses so that learners were encouraged in their efforts. In their discussion of how people learnt, Bransford et al. (1999) argued that not only did people come to formal learning with pre-existing knowledge, they also possess the capacity that enabled them to access new knowledge through active learning and metacognition. Metacognition referred to knowledge about learning, (a) about one's own learning, (b) about personal strengths and weaknesses, and (c) self-regulation through recognition of what was understood and when more information was required. Metacognition grew gradually with knowledge and experience. Although self-regulation might appear quite early, the ability to reflect on one's acquisition of knowledge would seem to develop at a later stage. Bransford et al. (1999) suggested that teachers could assist children to reflect on their work through activities such as working with other children on a topic; by reviewing the work of other children; and by positive reinforcement.

Sperber and Wilson (1995, p. 3) described the message contained in communication systems as "a code which pairs messages with signals". They considered this code model to be relevant to all forms of human communication. However, they argued that the code model of linguistic communication did not adequately account for the gap that might occur between semantic representation of the sentence and the thought that it was intended to convey. Sentences might be

ambiguous due to possible influences on the spoken word, such as where a statement might be altered to a question by the upward inflexion of the voice at the end of the utterance. Similarly, the interpretation of the sentence may be affected by other circumstances such as the status of the speaker, the context, or the purpose. Sperber and Wilson described this as the inferential dimension of communication and considered inference as an important aspect of the decoding process.

Consequently, Sperber and Wilson (1995) argued, although individuals might share a language they might not have the same assumptions about the world and the only way to avoid misunderstanding would be to ensure that the contexts of speaker and hearer were identical. Moreover, ambiguity frequently occurred in speaking. Therefore, in order to access the correct message, the recipient needed to decode both the semantic structure of the utterance and the inferential information encoded therein. This would be achieved through choosing the most contextually relevant meaning. Sperber and Wilson (1995) identified two forms of assumption presented in utterances. They were the explicature or explicit communication, and the implicature or implied assumption which was a combination of linguistically encoded and contextually inferred conceptual features. It might be extrapolated that, wherever possible, teachers should use explicature type utterances in order to avoid ambiguity.

Non-verbal Communication

Non-verbal communication might occur in association with spoken language or independently. Kendon (1981, p. 3) described non-verbal communication as:

[the] communicational functioning of bodily activity, gesture, facial expression and orientation, posture and spacing, touch and smell, and of those aspects of utterance that can be considered apart from the referential content of what is said.

Human verbal utterances may be understood in a variety of ways because in most cases more than one semiotic system is being invoked at any time. For example, the words in a sentence, spoken or written, are influenced by the context and the purpose of the communication or encounter, so that a constructed statement may inform, question, order, or contain subversive information known only to the participants. Utterances may be accompanied by gestures such as wink, fist thumping, and shoulder shrugging. Voice intonation may occur as a whisper, a shout, or be terminally raised. Proxemics includes touching, averting, and standing from being seated.

What is important in the interaction between teacher and learner, particularly the young learner new to formal education, is the disparate status of the child and the teacher. Givens (1981), referred to a study of three to five year olds who, when approached by unknown adults, averted their gaze, and whose heartbeat rates increased accordingly. These examples indicated that gaze aversion was one way that children could control their socially induced stress. Therefore the possibility existed that in the uneven levels of status of teacher-student communication, direct gaze with eye contact might be interpreted as challenging or threatening to the child.

Since gaze is such a forceful form of communication, it is more productive to the learning process for teachers to neutralise this activity as far as possible. An effective strategy for managing gaze was demonstrated by Teacher B who avoided lifting her gaze to the level of the child's face. Instead, she focused on a point which could be shared with the child such as the book being read, or piece of work being undertaken by the child. In their classrooms both Teacher A and Teacher B maintained the height of their seating arrangements consistent with that of the children in any situation where discussion was undertaken. This avoided any implication that the teacher, redolent with authority, was looking down on the children, and enabled the teacher to establish a more egalitarian context in student-teacher interaction.

Sperber and Wilson (1995) described style as the relationship between a speaker and his or her assumptions about the listener. It is a response to what the speaker considers to be the cognitive capacity and attention span of the hearers, and how much guidance is required by the hearers. Teachers constantly monitor their students and since they continually adapt their utterances to suit their classes, they select the most effective style of communication.

SUMMARY

The five topics that grouped the identified successful teaching processes and practices for Australian multicultural classrooms are reorganised in this article into two student-focused themes for discussion. The two themes are: the learning environment and student-teacher communication. The theme of the learning environment describes scaffolding and transfer as methods of instruction; holistic learning as a method of knowledge organisation and interpretation, including values education; and Gardner's theory of multiple intelligences as a strategy to address the individual talents and abilities of students. The theme of communication is sub-divided into verbal and non-verbal communication. Verbal communication includes: a) prior-learned language; (b) language as a semiotic system; (c) questioning; (d) feedback; and (e) inference and ambiguity. Non-verbal communication includes: a) gesture; (b) voice intonation, (c) proxemics and gaze.

The identified successful teaching processes and practices are discussed in this article in the context of relevant educational, psychological and anthropological research and discourse with relevant examples of successful teaching strategies. This discussion constitutes a guide for teaching in an Australian multicultural classroom. These identified successful pedagogical processes and practices recognise, acknowledge, and value the cultural diversity of an Australian classroom. In doing so, they support each child's endeavour to meet prescribed syllabus outcomes by developing his or her innate potential and by building on early learning. These strategies prepare all children for lifelong learning by providing them with basic skills.

REFERENCES

- Abercrombie, D. (1972). Paralanguage. In J. Laver and S. Hutcheson (Eds.), *Communication in Face to Face Interaction*. London: Penguin Books.
- Adams, D. (1968). The monkey and the fish: Cultural pitfalls of an educational adviser. In A. Dundes (Ed.) *Every Man his Way – Readings in Cultural Anthropology*. Englewood Cliffs, NJ: Prentice-Hall.
- Alix, N. M. (2000). The theory of multiple intelligences: A case of missing cognitive matter. *Australian Journal of Education* 44 (3):272-293.
- Armstrong, T. (2000). *Multiple Intelligences in the Classroom*. (2nd edition). Alexandria, VA.: Association for Supervision and Curriculum Development.
- Aspin, D., Chapman, J. and Klenowski, V. (2001.). Changing cultures in schools in Australia. In J. Cairns, D. Lawton and R. Gardner (Eds.) *Values, Culture and Education, World Year Book of Education 2001*. . London: Kogan Page Limited.
- Bransford, J. D. Brown, A. L. and Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.
- Brown, R. (1974). The development of the human child's native language. In A. Silverstein (Ed.), *Human Communication: Theoretical Explanations*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Delors, J. (1996). *Learning: the Treasure Within*. Paris: UNESCO.
- Duffy, D. (1994). Holistic Thinking in education in Australia. In D. Duffy and H. Duffy (Eds.) *Holistic Education: Some Australian Explorations*. Belconnen, ACT: Australian Curriculum Studies Association (ACSA).
- Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

- Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*. New York: Basic Books.
- Garrett, P. B. and Baquedano-Lopez, P. (2002). Language socialization: Reproduction and continuity, transformation and change. *Annual Review of Anthropology* 31:339-361.
- Geertz, C. (1973). *The Interpretation of Cultures - Selected Essays*. New York: Basic Books.
- Givens, D. (1981). Greeting a stranger. In A. Kendon (Ed.) *Nonverbal communication, interaction and gesture: Selections from Semiotica, Approaches to Semiotics*. The Hague: Mouton.
- Hall, E. T. (1976). *Beyond Culture*. New York: Anchor Press-Doubleday.
- Hansen, J. F. (1979). Education as Process. In J. Hansen (Ed.) *Sociocultural Perspectives on Human Learning: An Introduction to Educational Anthropology*. Eaglewood Cliffs, NJ: Prentice-Hall.
- Hawkes, T. (1977). *Structuralism and Semiotics. New accents*. London: Methuen.
- Heath, S. B. (1982). Questioning at home and at school: A comparative study. In G. Spindler (Ed.) *Doing the Ethnography of Schooling: Educational Anthropology in Action*. New York: Holt, Rhinehart and Winston.
- Kendon, A. (1981). Introduction. In A. Kendon (Ed.) *Nonverbal Communication, Interaction, and Gesture, Selections from Semiotica, Approaches to Semiotics*. (pp. 1-53). The Hague: Mouton.
- NSW DET. (2002). *K-YR 6 Assessment & Reporting 2002*: New South Wales Department of Education and Training. [Online]. Available: <http://www.bos.nswk6.nsw.edu.au-learning-k6assessments-index.php>.
- Schieffelin, B. B. and Ochs, E., (Eds.) (1986). *Language Socialization Across Cultures*. Cambridge: Cambridge University Press.
- Sperber, D. and Wilson, D. (1995). *Relevance - Communication Cognition, (2nd edition)*. Malden, MA: Blackwell.
- Teasdale, G.R. (1999). Implication of the Delors report for schooling in South Australia. In J. P. Keeves (Chair) *Invitational Seminar on the Delors Report: Learning - The Treasure Within*, (pp. 76-83). Adelaide, Australia: Flinders University Institute of International Education.
- Teasdale, G. R. and Teasdale, J. I. (1999). Culture and schooling in aboriginal Australia. In E. Thomas (Ed.), *International Perspectives on Culture and Schooling - A Symposium Proceedings*. London: University of London.
- Winch-Dummett, C. A. (2004). Learning to live together in an Australian multicultural classroom: A response to Delors. Unpublished PhD thesis. Adelaide, South Australia: Flinders University.