Science and Mathematics Education Centre

Influence of Students' Prior Knowledge and Classroom Learning Environment on Self-Efficacy and Achievement in Mathematics

Venkata Lakshmi Narayana Aluri

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

March 2018

DECLARATION

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number # SMEC 66-14

Signature: .....

Venkata Lakshmi Narayana Aluri

5 March 2018

ACKNOWLEDGEMENTS

I wish to take this opportunity to express my sincere gratitude to my Supervisor, Professor Barry J Fraser. I am greatly indebted to him for all the guidance, constructive comments and intellectual inspirations that made me complete this research within the prescribed time. His dedication, encouragement and support motivated me to strive and do my best in this academic journey. I extend my sincere thanks to Associate Professor Jill Aldridge, my Co-supervisor, for all the valuable suggestions, directions and support. My supervisors' active involvement and interest in my work motivated me to present part of this research at conferences. I would like to extend my sincere thanks to Professor David Treagust, who has helped me as a student and researcher. My appreciation also goes to Ms Petrina Beeton and Mrs Rosalie Wood at the Science and Mathematics Education Centre, as well as the staff of the Graduate Research School of Curtin University for their administrative support.

The presentation of my research at the participating schools in India and Australia has been rewarding learning experience. My special thanks to the school managements and principals for their permission, as well as teachers and students for their participation in my surveys in India. My special thanks also go to the Department of Education and Early Childhood Development, State Government of Victoria, who kindly allowed me to conduct my research survey in a secondary school in Victoria.

This piece of work could not have been completed without the unconditional love of my family members, especially the ever-smiling inspirations and rechargeable energies I have gained from my beloved granddaughters, Niharika and Madhu who made their debut as mathematics learners under my care within their zone of proximal development. I am indebted to my first daughter, Anjana and her husband, Venugopal Boyanapalli for their thought-provoking ideas during the developmental phases of this thesis, as well as my second daughter, Priyanka for her encouragement in my endeavour. The final heartfelt thank-you goes to my wife, Mani for her invaluable encouragement, patience with her gift of love. Mani, you have always provided me unlimited support with caring love and inspiration throughout my arduous research journey.

ABSTRACT

The lack of research on the linkages between students' classroom learning environment, prior knowledge, self-efficacy beliefs and achievement, particularly in India and Australia, provided the impetus for this research. This study's main aim was to investigate the influence of prior knowledge and classroom learning environment on self-efficacy and achievement in mathematics.

Data were collected from 464 Year 10 (or Class X in India) mathematics students and their 11 teachers from 15 classes of an Australian school and five Indian schools using: a) a knowledge framework to design a pretest of students' prior knowledge; b) a teachers' reflective journal for recording how teachers used their students' prior knowledge in classroom teaching; c) a 56-item, seven-dimensional Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) for seeking students' perceptions of classroom learning environment; and d) a 31-item achievement posttest and corresponding Mathematics Self-Efficacy Scale (MSES) for seeking students' self-reports of efficacy judgements in successfully achieving tasks on the posttest. My study has five research questions.

The first research question involved how teachers used their students' prior knowledge for teaching. A qualitative analysis of feedback from reflective journals revealed that many teachers introduced the new topic of measurement using three-dimensional objects in whole-class teaching, divided students into small groups using their prior knowledge assessment, motivated low-ability students by recalling prior knowledge with enabling prompts, assigned them challenging open-ended tasks and reinforced understanding with more classroom practice and homework.

The second research question focused on students' prior knowledge and their achievement of a topic of measurement. A quantitative analysis identified pedagogical approaches that offer equally-challenging opportunities for students of all ability levels and improve their achievement. Significant gender differences were found in achievement but not in prior knowledge for the whole sample.

The third research question required measurement of students' perceptions of classroom learning environment. I developed a 56-item seven-scale Mathematics-

related Constructivist Oriented Learning Environment Survey (MCOLES) after modifying the pre-validated 88-item COLES used in science. Analysis of MCOLES data resulted in distinct contributions to the field of learning environment: a) extracting ten MCOLES factors by exploratory and confirmatory factor analyses; b) identifying the relative importance of different MCOLES factors by analysing correlations with achievement; and c) a single-factor representation of the construct of classroom learning environment by second-order confirmatory factor analysis, because first order-factors were cross-correlated, theoretically and empirically. This enabled investigation of gender differences in partial correlations between students' achievement and classroom learning environment as a latent variable, with partial correlations being significant for boys but not for girls in some schools surveyed due to the confounding influence of covariables on achievement.

The fourth research question involved developing a classroom tool for calibrating students' self-efficacy judgements into efficacy-expectancies (expected scores on an achievement posttest). This tool proved useful for identifying that: a) gender differences in efficacy-expectancies were significant for some schools; b) the extent of underachievement due to inadequate prior knowledge was about 8% only, and due to over-assumed capabilities with reasonable prior knowledge was about 63%. Thus, this tool was helpful in identifying underachievers needing remedial counselling and an appropriate teacher intervention.

For the fifth research question, I investigated hierarchical data by a variance component analysis for the maximum likelihood estimation of the mediation model and a Two-level structural equation model (SEM) of students' achievement. The direct and mediated effects of classroom learning environment on achievement were statistically significant for high-ability students but not for low- and medium-ability groups. The direct and mediated effects of prior knowledge on achievement were statistically significant for the whole sample, while being higher for the low- and medium-ability groups combined than for the high-ability group. The joint influences of classroom learning environment and prior knowledge on achievement, estimated from the SEM, were relatively smaller than the corresponding direct effects due to the covariance between the determinants of achievement.

TABLE OF CONTENTS

page

Page

Page

Page

Page

page

LIST OF TABLES

page

page

page

LIST OF FIGURES

page

LIST OF APPENDICES

Chapter 1

INTRODUCTION

1.1 Introduction

The aims of my study were to investigate how teachers use their students' prior knowledge for teaching mathematics in their classes, and the extent to which students' prior knowledge and mathematics classroom learning environment jointly influence their self-efficacy beliefs and achievement. For accomplishing this, I considered a mathematics topic of measurement at the Year 10 level of the curriculum and used a pooled sample of 531 students and their 11 mathematics teachers recruited on a voluntary basis from Australian and five Indian schools.

This chapter has six sections. Section 1.2 briefly presents a global perspective of students' achievement based on results of the Programme for International Student Assessment (PISA), including a rationale for selecting Year 10 students from schools in Australia and India. Section 1.3 provides a preliminary discussion of prior knowledge and classroom learning environment as determinants of self-efficacy and achievement. Later, Section 1.4 briefly outlines theories surrounding the relationships posited in my research model. Section 1.5 provides a list of my research questions and research objectives, as well as research methods, instruments, data analyses and limitations of my sample. Section 1.6 explains the significance of this research. Section 1.7 presents the organisation of my thesis.

1.2 Students' Achievement in Mathematics at the Year 10 Level

As an indicator of successful teaching and learning, students' achievement is often used to improve the quality of teaching. Reviews of OECD studies in mathematics and science education (OECD, 2016a) cautioned about the decline in student achievement and enrolments in mathematics courses and pointed to an

immediate need for curriculum changes focused on engaging all young people in mathematics in OECD countries (Niemann, Martens & Teltemann, 2017). For Australia, previous results of PISA indicated a steady decline in performance, with mean scores in mathematics for Australia falling from an estimated average of 524 points in 2003, to 514 in 2009, to 504 in 2012 and to 494 by 2015, which is far below the leading performance of 564 points displayed by Singapore (Gurria, 2016).

Stacey and Turner (2015) commented that, in many countries, PISA results of mathematical literacy, with its analysis of what makes mathematics education useful for most future citizens, have been extremely influential in curriculum review and for improving teaching and learning. Whilst country-specific curricula revisions are important, these reviews provide limited evidence of factors that explain the relatively low level of achievement at the Year 10 level.

The rationale for my research focus on students' achievement at the Year 10 level in Australia and India stems from three reasons. First, students at the Year 10 level are in transition from secondary to senior secondary school which is when students are expected to focus on performance-oriented academic achievement and making peer comparisons, and when their task interest matters in their motivational self-efficacy beliefs in achieving. For instance, Cleary and Chen (2009, p. 1) reported that "task interest was shown to be the primary motivational predictor …during math learning".

Second, my study used a pooled sample of mathematics students and teachers from India and Australia because the bilateral agreements between these countries (see reports of Australia-India BRIDGE School Partnership program from www.australiaindiaeducation.com) facilitated continuous migration flows of student populations from India to Australia for secondary and higher education. As noted by Rizvi and Gorur (2011), academic links between Australia and India have a long history, stretching back to the late nineteenth century, but blossoming since the 1950s. Under the Colombo Plan, Australian universities provided India with wide-ranging development assistance in education, with a focus on skills formation in science and technology, and many Indians were awarded scholarships to undertake advanced studies at Australian universities. The late 1980s witnessed an increasing number of fee-paying Indian students attending Australian education and training institutions. In

2010, more than 100,000 students from India enrolled in Australian institutions of higher and vocational education (Rizvi, Gorur & Reyes, 2015).

Also, the Australia India Education Council (AIEC) identified new forms of social and economic coordination and significant synergies, which could be harnessed if both countries are to mutually profit in a meaningful and sustained manner through collaborations in higher education (Rizvi, Gorur & Reyes, 2015).

Third, for India, there is a limited evidence about students' achievement at the Year 10 level (or Class X) because India did not participate in the PISA (Chhapia, 2012). Therefore, a national assessment survey of Class X was undertaken by the National Council of Education Research and Training (NCERT) in 2014-2015, which revealed low-to-average achievement by most students across different States in India, but not their causes (Sreekanth et al., 2015).

My study, therefore, sought to fill a research gap and endeavoured to provide an understanding of research-based cognitive and non-cognitive determinants of students' achievement, and quantify their effects to inform appropriate classroom actions to improve students' achievement in future.

The following sections offer an overview of pertinent determinants of students' achievement, as well as the research questions and objectives addressed in this study.

1.3 Determinants of Students' Mathematics Achievement

To investigate the sources of variation in students' mathematics achievement, I conducted the detailed literature review reported in Chapter 2. My review identified factors influencing achievement, including important issues of diversity for mathematics teaching and learning, such as: students' prior knowledge required for mathematics learning, classroom teaching practices of tracking or streaming (ability grouping); gender differences in prior knowledge and achievement; and classroom learning environment and self-efficacy beliefs in achieving given tasks. These are explained in that order in the following subsections.

1.3.1 Students' Prior Knowledge

Mathematics is a language that is socially constructed in classrooms, and it is learnt best when taught in a constructive way relating students' background and prior knowledge to learning by using suitable mathematics tasks. This was substantiated by Ernest's (1991, 1998) social philosophy of mathematics learning, von Glasersfeld's (1985, 2000) constructivists theory and Dochy's (1992, 1994) seminal works on prior knowledge for learning. Past studies about learning theories lend support for prior knowledge as a key construct and predictor of students' achievement (Ausubel, 1968, 2000; Dochy, 1992; Schneider, Korkel, & Weinert, 1989; Tobias, 1994).

For operationalising prior knowledge, many cognitivists recognised the notion of conditional knowledge and two types of declarative knowledge and procedural knowledge (Alexander, Schallert, & Hare, 1991; Paris, Lipson, & Wixson, 1983). Following this, Hailikari, Nevgi, and Lindblom-Ylanne (2007) provided an operational framework, as shown later in Table 2.1 in Subsection 2.3.2, that describes the nature of mathematics tasks, as a guide for designing mathematics tasks to assess students' knowledge.

By assessing prior knowledge according to this framework, many researchers investigated plausible relationships between students' prior knowledge and their learning outcomes and acknowledged a positive significant influence of prior knowledge on learning (Bloom, 1976; Dochy, 1994, Weinert & Helmke, 1998). Furthermore, Hailikari et al. (2008, p. 59) examined "relationships between prior knowledge, academic self-beliefs and previous study success in predicting achievement of college majors", and reported that, "domain-specific prior knowledge was the strongest predictor of student achievement over and above other variables included in their model".

1.3.2 Ability Grouping

In a mixed-ability classroom setting, students' achievement could be impacted by classroom practices adopted for teaching, such as ability grouping. Past research has focused on classroom teaching and assessment practices for two different forms of

grouping: i) homogeneous grouping or streaming, when students are assessed in mathematics and then grouped according to the results on that assessment and ii) heterogeneous grouping or tracking, when students are provided with opportunities for accelerated learning by selecting one or more high-achieving groups, while the rest of the students are grouped heterogeneously (Sullivan, 2015).

There is limited research on the extent of streaming practices in Victorian schools. In a detailed on-line survey-based study, Forgasz (2010a, b) reported that 80% of the 44 schools responding had some form of streaming in the years 7-10, with three quarters of the respondents indicating support for that streaming.

Marsh et al. (2005) and Sullivan (2015) recognised that both methods of ability grouping are fraught with problems of equity in learning, as described later in Subsection 2.2.2. Sullivan (2015) recommends that all students in mixed-ability class settings should have ready access to challenging learning opportunities in the interest of equity. Askew (2015) recommends that diversity can be reduced by managing differences between learners, such as by promoting teaching practices that encourage individualised learning experiences in a mixed-ability classroom setting.

Internationally, a meta-analysis by Steenbergen, Makel and Olszewski (2016) synthesised approximately 100 years of research in the United States on the effects of ability grouping and acceleration on K-12 students' academic achievement. Outcomes of 13 meta-analyses of ability grouping showed that students benefited from WITHIN Class grouping (mixed-ability classroom setting without tracking), with effect sizes ranging between 0.19 standard deviations (SD) and 0.30 SD, but not from BETWEEN Class grouping (0.04-0.06 SD) (as in streaming). The effects did not vary for high-, medium- and low-ability students. For Indian classrooms, there is little evidence on ability grouping practices and their effects on students' achievement.

Thus, to fill this research gap, I attempted to examine the achievement of students of each ability group, after teachers adopted WITHIN Class grouping without tracking or streaming and based on a pretest of prior knowledge of a topic, which was designed according to the Hailikari, Nevgi, and Lindblom-Ylänne's (2007) knowledge framework. At the same time, as recommended by Sullivan (2015), teachers in my study taught a topic of measurement by a judicious selection

of mathematics tasks drawn from the prescribed curriculum and provided equally-challenging learning opportunities for students of all ability groups.

1.3.3 Mathematics Classroom Learning Environment

Fraser's (2007, 2014) past research over the past 40 years has consistently shown that the quality of classroom learning environment in schools in Australia and other countries is an important determinant of students' learning and that students are likely to learn better when they perceive a positive classroom environment.

Past research recognised the constructivist approach to mathematics learning in developing items for seeking students' perceptions of classroom learning environment in different countries (Aldridge, Taylor, & Chen, 2000; Fraser, 2002, 2012; Taylor, Fraser, & Fisher, 1997) and, thus, the Constructivist Learning Environment Survey (CLES) (Taylor et al., 1995; Taylor et al., 1997) was developed, used in science learning, and extensively adopted in different countries.

Following Fraser and Tobin (1991) and Tobin and Fraser (1998), Taylor, Fraser and Fisher (1997) and Aldridge et al. (2000) modified the original version of the CLES by omitting negative items and arranging them systematically to give a cue to the respondents, rather than randomly as in many previous questionnaires. Aldridge et al. (2000) designed the 30-item CLES with five key dimensions of critical constructivist learning environment, introduced a five-point response format of Almost Always, Often, Sometimes, Seldom and Almost Never and validated CLES by administering it to 1,081 students from 50 classes in Australia and 1,879 students from 50 classes in Taiwan. For India, there is a limited evidence of studies that adopted CLES.

In a subsequent study, Aldridge et al. (2012) incorporated additional dimensions of Personal Relevance, Shared Control and Uncertainty in developing the Constructivist-Oriented Learning Environment Survey (COLES) that has 88 items spread across 11 dimensions. COLES dimensions also match the Six Principles of Learning and Teaching followed in Victorian schools (State Government of Victoria, 2012). Aldridge et al. (2012) validated the COLES and examined associations between classroom learning environment and learning outcomes, including students' attitudes and achievement. However, few studies applied COLES to the mathematics

domain. My study filled this research gap after modifying and reducing the 88-item COLES to form the more-efficient 56-item Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) that requires less response time.

The following sub section deals with the role of self-efficacy beliefs in achieving tasks, as a determinant of achievement.

1.3.4 Self-Efficacy Beliefs in Achieving Mathematics Tasks

Bandura's (1986, p. 22) social cognitive theory construes "human functioning as a series of reciprocal interactions between personal influences, environmental features and behaviours". Bandura (1986a, p. 25) claims that "what people think, believe, and feel affects how they behave". This supports key relationships between students' personal thoughts and beliefs, and classroom environment and behaviours including students' focussed attention and effort in learning.

Schunk and Pajares (2004, p. 115) noted that "...in educational settings, self-efficacy is not an isolated construct but rather an integral component of social-cognitive theory, because students' social environment can influence both their affective domains and behaviours, as evident in classrooms".

Past research emphasised that students' successful learning in mathematics is mainly determined by both their sustained motivation for mathematics learning and their motivational beliefs in achieving given mathematics tasks. Various researchers have investigated the role of self-efficacy beliefs in students' mathematics classroom learning environment and achievement (Collins, 1982; Fast, Lewis & Bryant, 2010; Schwarzer, 2014)

Velayutham (2012), in her study of the influence of psychosocial environments on students' motivational beliefs of self-efficacy, noted that classroom learning environment research focuses on classroom life, usually from the students' perspective.

Dorman (2001) and Dorman and Adams (2004) reported positive and statistically-significant correlations between academic self-efficacy and ten classrooms learning environment dimensions. Relationships with self-efficacy effects, however, were statistically nonsignificant for Student Cohesiveness, Teacher Support, Personal Relevance and Shared Control.

Hailikari, Nevgi, and Komulainen (2007, p. 59) defined academic self-beliefs as “an individual’s beliefs about his or her attributes and abilities as a learner”. By measuring students’ academic self-beliefs to a set of items that they developed, Hailikari et al. concluded that “academic self-beliefs strongly correlated with previous study success, had a strong direct influence on prior knowledge test performance, and predicted student achievement only indirectly via prior knowledge. These results imply that both prior knowledge and self-beliefs should be considered when considering instructional support issues because they can provide valuable insights about the future performance”.

But Bandura (1980) made a strong case for considering self-efficacy at a task-specific level. That is, self-efficacy should be investigated for individual tasks rather than using global indicators such as confidence in learning index (Fennema & Sherman, 1976), because of the behavioural and situational specificity of self-efficacy.

Following this, Hackett and Betz (1989) defined perceived self-efficacy as a situational capability to achieve a task. Hackett and Betz (1989) designed mathematics tasks and developed a self-efficacy instrument for seeking students’ perceptions of self-efficacy beliefs in achieving given tasks successfully. They calibrated students’ self-efficacy judgements into efficacy-expectancies (expected scores) as a measure of self-efficacy beliefs (Bandura, 1980, 1986, 2006).

Further, students’ achievement scores as awarded by their teachers were compared with students’ expected scores to identify underachievers and the extent of underachievement, as well as those students with expected scores that were lower than what they achieved. Bandura (1980) emphasised that, given incentives to perform, efficacy-expectancies are a major determinant of people’s choice of activities, how much effort they expend, and how long they sustain effort in dealing with stressful situations.

Furthermore, many researchers have investigated the mediational role of self-efficacy beliefs and quantified the mediated influence of student-related cognitive factors including prior knowledge on achievement (Fast et al., 2010; Hackett & Betz, 1989; Hailikari, Nevgi, & Komulainen, 2007; Randhawa et al., 1993; Siegel et al., 1985).

1.3.5 Rationale for Investigating Gender Differences

The rationale for investigating gender differences in my study stems from contemporary mathematics education reviews which suggest that gender differences continue to exist for student achievement and the selection of STEM courses and careers (Fennema & Leder, 1990; Forgasz & Leder, 2015, 2017; Murphy & Gipps, 1996; Parker, Rennie, & Fraser, 1995). Recently, Tellhed, Bäckström and Björklund (2016, p. 86) argued: "Throughout the world, the labor market is clearly gender segregated. More research is needed to explain women's lower interest in STEM (Science, Technology, Engineering and Mathematics) majors and particularly to explain men's lower interest in HEED (Health care, Elementary Education, and the Domestic spheres) majors." A similar argument was echoed by Forgasz and Leder (2017) when they compared the general public's perceptions with 15-year-old school students' attitudes towards mathematics learning from the large-scale PISA (2012) and found that, "…while participants considered mathematics to be important for everyone to study, and important for employment, vestiges of traditional gender stereotyped beliefs and expectations were evident, more so among the younger than older respondents" (p. 261).

Also, as mentioned before, past studies have highlighted the important role of students' self-efficacy beliefs as motivational determinants (Schunk, 2012) that affect their confidence, approach/avoidance of an academic task, persistence and performance in academic domains (Bandura, 1986b, 1997). For example, Velayutham, Aldridge and Fraser (2012) investigated gender differences in student motivation in science learning.

Kruger (1999) examined gender differences in self-efficacy perceptions and patterns of science achievement among 154 Year 9 boys and girls on multiple-choice and constructed-response items. Similarly, in the present context of mathematics learning, gender differences in self-efficacy beliefs could explain variations in achievement, Moreover, it is vital for mathematics educators to know of and respond to differences in the ways in which boys and girls learn mathematics at the year 10 level, as well as to investigate gender differences in students' perceptions of classroom learning environment.

Reviews by Forgasz and Leder (2015, 2017) pointed to the persistence of traditional stereotypes of mathematics as a male domain in Australia. In line with this challenge, the present study explored the role of gender differences in students' perceptions of classroom learning environment, prior knowledge required for mathematics learning, self-efficacy beliefs in achieving assigned tasks, and achievement of mathematics tasks at the year 10 level. The investigation of gender differences by ability groups based on students' prior knowledge could throw light on other aspects of this problem. Compared with familiar methods applied in most other gender studies, my research methods were different in that I examined gender differences in prior knowledge and achievement and in the correlations between prior knowledge and achievement.

1.4 Research Model and Theoretical Rationale

This section presents the research model, the underpinning theoretical rationale and the research objectives of this study.

1.4.1 Research Model

This study's proposed research model is shown in Figure 1.1 with two explanatory variables: students' prior knowledge required for learning; and the construct of mathematics classroom learning environment. The construct of students' self-efficacy beliefs in achieving given mathematics tasks acts as the mediator of effects of the explanatory variables on the criterion variable of students' achievement in mathematics. These relationships are underpinned by the theoretical rationale explained below.

1.4.2 Theoretical Rationale

The research model and hypothesises of my study were based on theories surrounding mathematics learning such as constructivist theory (von Glasersfeld, 1985, 2000), philosophy of social constructivism (Ernest, 1991, 1998), social-cognitive theory and self-efficacy theory involving classroom learning environment. Addressing the centrality of the self-efficacy mechanism in human agency, Bandura (1982, p. 122) states that "self-precepts of efficacy influence thought patterns, actions, and emotional arousal."

In the field of education, Urdan and Schoenfelder (2006, pp. 340-341) argue that "schools and classrooms are, by definition, social environments. Teachers can make their students responsible learners by enhancing their positive self-beliefs and paying attention to the key features of classroom learning environment.

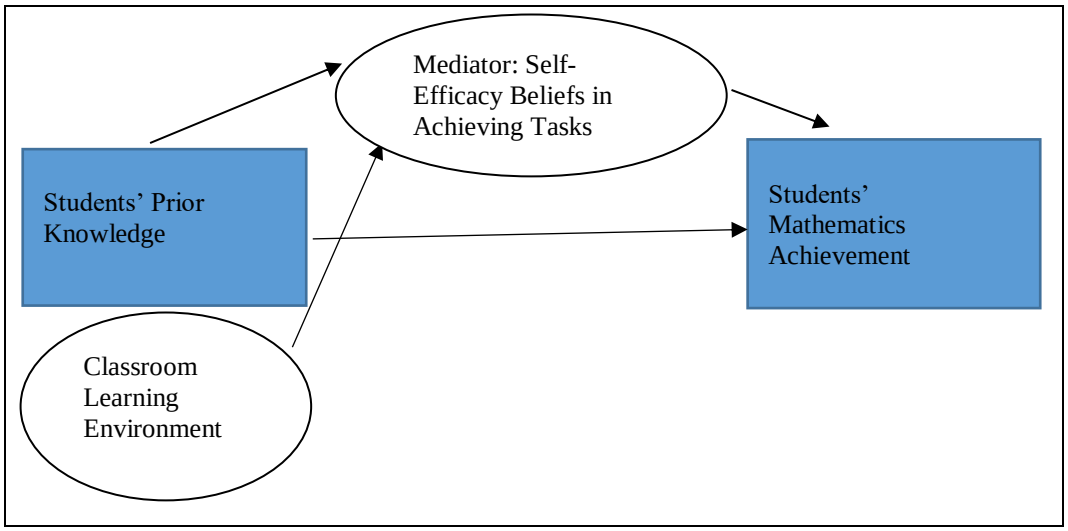


Figure 1.1: A Schematic View of the Research Model Investigated in this Study

Underpinned by the above theoretical relationships, the present study included psychosocial aspects of learning environment in the research model for quantifying their influence on students' self-efficacy beliefs in mathematics achievement.

1.5 Research Questions, Objectives, Methods, Instruments and Data Sources

This section has four subsections. It presents research questions in Subsection 1.5.1, research objectives in Subsection 1.5.2, research methods, instruments and data sources in Subsection 1.5.3, and limitations of the sample in Subsection 1.5.4.

1.5.1 Research Questions

Broadly, five research questions were identified from the above discussion of hypothesised relationships involving the influence of students' prior knowledge and classroom learning environment on students' self-efficacy beliefs and achievement, as well as gender differences in prior knowledge, in self-efficacy expectancies in achieving, and in mathematics achievement of students of different ability groups:

1. How do teachers use their students' prior knowledge when teaching a topic of the year 10 mathematics curriculum?
2. a) Are there gender differences in students' prior knowledge and achievement?
   b) Is students' achievement associated with i) prior knowledge required for learning a mathematics topic, and covariables of ii) teaching experience and iii) class size?
   c) Are there improvements in students' achievement when teachers select suitable tasks for a mathematics a topic, when teaching students of different ability groups?
3. a) Is an instrument to measure students' perceptions of mathematics classroom learning environment at the Year 10 level theoretically-sound and valid?
   b) Are students' perceptions of classroom learning environment associated with i) prior knowledge required for learning and ii) achievement in the chosen topic?
   c) Are there gender differences in correlations between students' perceptions of classroom learning environment and achievement?
4. a) Can a classroom tool be developed to identify underachievers and students with low self-efficacy in achieving, so that a suitable teacher

intervention program can be suggested to improve achievement?

b) Are students' self-efficacy expectancies (expected scores) on mathematics tasks set on a test associated with i) classroom learning environment ii) prior knowledge required for learning and iii) achievement?

c) Are there gender differences in correlations between students' self-efficacy expectancy and achievement?

5. Do students' prior knowledge and mathematics classroom learning environment jointly influence their achievement, as mediated by their self-efficacy expectancies?

1.5.2 Research Objectives

To answer the five research questions raised above, the following research objectives were identified for the present study:

1. To address the first research question, the research objectives were to design a pretest of prior knowledge required for learning a topic selected from the mathematics curriculum at the year 10 level and administer it to students in consultation with their teachers, as well as to evaluate qualitatively how teachers utilised their students' prior knowledge when teaching that topic in mixed-ability classrooms.
2. To address the second research question, the research objectives were to design a posttest of achievement for the topic taught, administer it to students in consultation with teachers and use students' scores to examine:

   a) gender differences in students' prior knowledge and achievement for i) the whole sample and ii) in each school surveyed

   b) correlations between students' achievement in that topic and i) their prior knowledge required for learning that topic and ii) classroom covariables such as teaching experience and class size for i) the whole sample, ii) each school and iii) each class

c) improvements in students' learning relative to their prior knowledge for i) the whole sample, ii) each school, iii) each class and iv) each ability group in selected classes.

3) To address the third research question, this study's research objectives were to:

a) develop and validate the new Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES)

b) extract factor values for the construct of classroom learning environment by exploratory and confirmatory factor analyses

c) correlate classroom learning environment dimensions (factor values) with students' achievement and prior knowledge

d) examine gender differences in correlations between students' perceptions of classroom learning environment (factor values) and their achievement.

4) For answering the fourth research question, this study's research objectives were to:

a. develop the new Mathematics Self-Efficacy Scale (MSES) using the tasks set on the achievement posttest

b. use the MSES to calibrate students' self-efficacy judgements and estimate students' expected scores as a measure of their self-efficacy expectancies and, hence, to identify for teacher intervention underachievers and students with self-efficacy expectancies that were lower than their achievement

c. correlate students' self-efficacy expectancy (expected scores) with i) students' prior knowledge, ii) classroom learning environment (factor values) and iii) students' achievement

d. examine gender differences in correlations between students' self-efficacy expectancy (expected scores) and achievement.

5) To address the fifth research question, the research objectives were to:

a. posit a mediational model to carry out a mediation analysis of self-efficacy beliefs to estimate direct and mediated effects of students' prior knowledge and classroom learning environment on achievement

b. quantify the joint influence of students' prior knowledge and classroom learning environment on achievement by estimating and evaluating a Two-level Structural Equation Model (SEM) of students' achievement.

1.5.3 Research Methods, Instruments and Data Sources

For data collection, five research instruments were designed and administered to a sample of students drawn from one Australian school and five Indian schools as detailed in Chapter 3. For addressing the first research objective, I developed: i) the knowledge framework (Appendix 1) recommended by Hailikari, Nevgi and Lindblom-Ylänne (2007) and ii) a pretest of tasks (see Appendix 2) related to the topic of measurement selected from the Year 10 mathematics curriculum, designed with appropriate tasks by using the knowledge framework given in Appendix 1, after consulting teachers who administered it to students in their respective classes. Class teachers scored pretest and shared prescores with the researcher. Furthermore, teachers were guided in identifying students in their respective classes and grouping them using prescores as:

0-40% for the low-ability group

40-60% for the medium-ability group and

60-100% for the high-ability group.

Later, teachers taught the topic of measurement by selecting tasks from the prescribed curriculum that were suitable for each ability group and recorded their classroom experiences in a Teachers' Reflection Journal (TRJ) (see Appendix 3).

For addressing research objectives 2(a) and 2(b), an achievement posttest was designed (see Appendix 4 for Class 1 and Appendix 6 for all other classes) after consulting teachers and by choosing tasks according to the knowledge framework recommended by Hailikari, Nevgi and Lindblom-Ylänne (2007). Teachers then administered posttest to all students to assess their learning. For addressing research objective 2(c), postscores and prescores were used to estimate the magnitude (effect sizes) and statistical significance of changes in learning.

For meeting research objective 3(a), I developed and validated the new Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) (see Appendix 8) and, for objective 3(b), I extracted factor values by exploratory and confirmatory factor analyses to represent the construct of classroom

learning environment. For objective 3(c), bivariate correlations between factor values of MCOLES dimensions and students' achievement were analysed and, for meeting objective 3(d), gender differences in correlations were tested using Fisher's *r*-to-*z* transformation (Kendall et al., 1973) on students' postscores.

For addressing research objective 4(a) and 4(b), I followed Bandura (1977, 2006) and developed the new 31-item Mathematics Self-Efficacy Scale (MSES) (see Appendix 5 for Class 1 and Appendix 7 for all other classes) and collected students' responses to the MSES items (which are also the same as the tasks designed in the posttest, as mentioned earlier). MSES responses are on a scale of 1 to 5 where 1 represents the expected score of 0% and 5 indicates the expected score of 100%. Thus, the method of calibration recommended by Bandura (2006) was used to convert students' responses to mathematics tasks in the posttest into equivalent scores that they expected according to their self-efficacy judgements. Later, students' expected scores were compared with the postscores awarded by their teacher according to a pre-informed marking scheme of evaluation. When students scored less than what they expected (after adjusting for a margin of error of 5%), they were treated as underachievers. On the other hand, when students expected much lower scores than what they achieved (after adjusting for a margin of error), they were treated as having low self-efficacy expectancies.

For addressing research objectives 4(c) and 4(d), I conducted an analysis of partial correlations between efficacy-expectancies (expected scores) and achievement by holding constant the other confounding covariables of prior knowledge and classroom learning environment (Fisher, 1924; Kendall et al., 1973). Also, gender differences were explored for students' prior knowledge (prescores), self-efficacy expectancy and achievement.

For addressing research objective 5(a), a mediation model suggested by MacKinnon et al. (2007) was adopted, and direct and mediated effects of prior knowledge and classroom learning environment on achievement were derived, using self-efficacy as the mediator. Finally, for accomplishing research objective 5(b), a Two-level SEM as explained in Subsection 3.7.3 was estimated by the method of maximum likelihood using the M*plus* software program (Muthen & Muthen, 2008) for

the whole sample and the high-ability group separately. For this purpose, use was made of hierarchical data collected for teaching experience, class size, prior knowledge (prescore), coefficient of variation in prescores, self-efficacy beliefs (expected scores) and achievement (postscores). For representing students' perceptions of classroom learning environment as a single variable, factor values were extracted parsimoniously by second-order confirmatory factor analysis of MCOLES responses.

1.5.4 Limitations of Sample

For administering the above instruments, a sample of 531 students and their 11 Year 10 mathematics teachers was drawn from 15 classes, which was somewhat restrictive when obtaining maximum likelihood estimates of parameters involved in the second level of SEM because of limited degrees of freedom. As a result, only one explanatory variable could be used at a time at Level 2 in the model for explaining the variance in students' achievement BETWEEN Classes, therefore giving rise to a low explanatory power for the model.

Second, after data clean-up, the sample size was reduced from 531 to 464 complete records, which included only 51 observations (student records) from the low-ability group and 160 from the medium-ability group; these sample sizes were inadequate for carrying out Two-level SEM analysis for these two groups separately. Thus, the analysis was confined to the whole sample (N=464) and the high-ability group (N=303) only.

The third limitation concerns drawing inferences based on the sample of students who took part in the survey on a voluntary basis. The findings of my study are based on a purposive sample of 464 students from five Indian schools and one Australian school, which disproportionately represents underlying student populations in these two countries at the Year 10 level; this limits inferring comparisons between India and Australia in their classroom teaching practices or students' learning.

However, given the objectives of the study, the pooled sample was purposive for making generalised inferences using the sample-based findings about the effects of cognitive and noncognitive factors on achievement and other analytical aspects such as:

i) development, validation and factor analysis of the new instrument, MCOLES

ii) hypotheses testing of gender differences in students' prior knowledge and achievement

iii) hypotheses testing for changes in learning

iv) direct and mediated effects of self-efficacy beliefs

v) path coefficients of Two-level SEM.

1.6 Significance of Research

This study bridged a research gap by examining the joint influence of prior knowledge and classroom learning environment, represented parsimoniously by a single factor, on students' self-efficacy beliefs and achievement in mathematics learning. The theoretical contribution adds to the literature on learning environment and educational psychology, and benefits teachers to improve students' achievement in mathematics learning, as well as provides further directions for future studies.

The first methodological contribution of this study is that the COLES was modified to form the MCOLES, which is suitable for mathematics classes. To establish the validity of the newly-developed instrument, I employed a comprehensive construct validity framework involving face, convergent, discriminant, concurrent and predictive validities. The MCOLES has 50 items only and was used for measuring students' perceptions of the construct of classroom learning environment. This method could be used by future researchers who wish to develop and validate new questionnaires.

The second methodological contribution was in the extraction of factors from MCOLES responses, as well as an application of second-order CFA to obtain a parsimonious representation of the construct by a single factor. Furthermore, the uni-factor representation was found to be analytically amenable for estimating the influence of classroom learning environment jointly with prior knowledge on achievement

Third, I also developed a classroom tool for calibrating students' self-efficacy judgements into efficacy-expectancies, measured by students' expected scores in the achievement posttest. It thus demonstrated how teachers could use a similar procedure for other topics of their choice for identifying underachievers and students who displayed lower self-efficacy in achieving. Thus, this tool can be used by teachers when considering appropriate classroom interventions for improving the future achievement of boys and girls, separately.

Fourth, major contributions emerged from estimates of gender differences in students' prior knowledge and achievement, as well as differences in correlations of students' achievement with classroom learning environment and self-efficacy expectancies, which add to the literature of mathematics education research in secondary schools The results of direct and mediated effects of students' prior knowledge and classroom learning environment on students' achievement and their joint influence estimated from Two-level SEM could be used in future studies to derive information related to the mediational role of self-efficacy.

Finally, the fifth contribution of this study was that researchers and teachers were provided with some convenient tools to help with organising students in mixed-ability classroom into low-, medium- and high ability groups based on their prior knowledge, and with adopting teaching practices that offer equally-challenging learning opportunities for students of all ability levels and genders. Teachers could use the information to refocus their pedagogical approaches, for example, by implementing and evaluating instructional strategies that have the potential to increase the achievement of both boys and girls.

1.7 Overview of Organisation of Thesis

The remainder of this thesis is organised as follows. Chapter 2 presents a review of the relevant literature, mostly based on a detailed online search of secondary data, covering the main themes of research objectives in five sections. Section 2.1 reviews studies involving students' mathematics achievement at the Year 10 level,

which was a criterion variable of this study, issues facing teachers when addressing diversity in students' ability levels for classroom teaching, and gender differences in achievement in Australia and India, with international comparisons. Section 2.3 reviews studies about theoretical underpinnings of students' prior knowledge as a determinant of achievement, and studies that provide a framework for operationalising prior knowledge so that its influence on achievement can be investigated. Section 2.4 reviews research related to mathematics learning environment, including constructivist theory and the philosophy of social constructivism for mathematics learning, as well as instruments for measuring students' perceptions of mathematics classroom learning environment. Section 2.5 offers a review of studies of social cognitive theory and its application to mathematics education in the secondary schools. It explains the use of mathematics self-efficacy in the theoretical development of this study's research model. Also, this section reviews studies of past questionnaires used for measuring the construct of self-efficacy beliefs in achieving a given set of mathematics tasks.

Chapter 3 deals with research paradigms, methods, data instruments and procedures of analysis adopted in this study. Section 3.2 provides research paradigms adopted, while Section 3.3 describes research methods. Section 3.4 provides details of instruments developed and employed for data collection, while Section 3.5 describes sample selection, data collection and sample demographics for the pilot study and the main study. Section 3.6 details the procedures for validating instruments, which involved a construct validity framework to establish content, face, convergent, discriminant, concurrent and predictive validities, as well as the factor structure for the construct of classroom learning environment. Section 3.7 then elucidates data-analysis procedures for testing gender difference and correlations between students' prior knowledge, classroom learning environment dimensions, self-efficacy beliefs and achievement. Finally, Section 3.8 presents procedures adopted to estimate the mediation model with direct and mediation effects of the explanatory variables, prior knowledge and classroom learning environment on achievement, as well as the procedures adopted for estimating the SEM parameters and model evaluation using fit indices.

Results of this study are organised in Chapters 4 to 7. The relevant chapters and subsections which addressed the research questions and corresponding objectives (as in Subsections 1.5.1 and 1.5.2) are listed in the overview Table 1.1. Chapter 4 provides results pertaining to the first and second research questions in five sections. Section 4.2 reports students' prior knowledge as measured by prescores and used by teachers for selecting students into three ability groups, as well as for my qualitative analysis of feedback information provided by teachers in Teachers' Reflective Journals. Section 4.3 deals with the second research question involving gender differences in students' prior knowledge and achievement, while Section 4.4 covers an analysis of correlations between students' achievement and their prior knowledge and the two covariables of teaching experience and class size. Section 4.5 uses frequency distributions of prescores and postscores mainly for investigating changes in students' achievement in i) the whole sample of students, ii) each school, iii) each

Table 1.1: Thesis Chapter Sections Where Findings are Reported for the Research Questions and Research Objectives

Res. Questions	Res. Objectives	Chapters	Sections / Subsections
1	1	4	4.2 and 4.6.1
2	2 (a)	4	4.3 and 4.6.2
	2 (b)	4	4.4 and 4.6.3
	2 (c)	4	4.5 and 4.6.4
3	3 (a)	5	5.2 and 5.5.1
	3 (b)	5	5.3.1 to 5.3.4 and 5.5.1
	3 (c)	5	5.4.1 and 5.5.2
	3 (d)	5	5.4.2 and 5.5.3
4	4 (a)	6	6.2 and 6.5.1
	4 (b)	6	6.3.1, 6.3.2 and 6.5.1
	4 (c)	6	6.4.1 and 6.5.2
	4 (d)	6	6.4.2 and 6.5.2
5	5 (a)	7	7.3 and 7.6.2
	5 (b)	7	7.4.2 and 7.6.3

class within a school and iv) ability groups in selected classes.

Chapter 5 deals with the development and validation of the new MCOLES for measuring classroom learning environment. It provides an answer to the third research question by addressing research objectives 3(a) to 3(d) of Subsection 1.5.2 in five sections. Section 5.2 addresses research objective 3(a) involving the development of

the MCOLES, including establishing content and face validities, as well as convergent, discriminant and concurrent validities. Section 5.3 addresses research objective 3(b) involving extracting MCOLES factors based on exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Section 5.4 answers research question 3(c) by establishing the predictive validity of the MCOLES with an analysis of correlations between students' achievement and different MCOLES factors, and reports gender differences in these correlations for each school and the whole sample.

Chapter 6 addresses research objectives 4(a) to 4(d) of Subsection 1.5.2 For research objective 4(a), Section 6.2 describes the development of the new MSES for the measurement of self-efficacy beliefs in achieving mathematics tasks. For research objective 4(b), Section 6.3 provides details of the calibration of self-efficacy judgements using MSES responses into students' expected scores, as a measure of self-efficacy expectancies in achieving tasks to help to identify underachievers and students with lower self-efficacy expectancies. For research objective 4(c), Section 6.4 deals with predictive validity of the MSES by an analysis of correlations between efficacy expectancies and students' i) prior knowledge, ii) classroom learning environment and c) achievement. To address research objective 4(d), Section 6.4 reports gender differences in efficacy-expectancies and in correlations with achievement.

Chapter 7 addresses research objectives 5(a) and 5 (b) of Subsection 1.5.2. Section 7.2 examines the rationale for using Two-level SEM modelling involving a variance component analysis of the dependent variables based on hierarchical data. Section 7.3 examines the mediational role of self-efficacy and reports estimates of direct and mediated effects of prior knowledge and classroom learning environment on achievement. The research model has three variants involving different types of explanatory variables. Section 7.4 presents the SEM of students' self-efficacy and achievement and its three variants (SEM1 - SEM3) involving different pairs of explanatory variables in the regressions. Section 7.4 also reports regression effects of the joint influence on achievement, as well as hypothesis testing of estimated effects, while Section 7.5 compares the fit of alternative models for explaining the variance in students' achievement (ACH).

Chapter 8 discusses results for each of the research objectives in terms of theoretical, methodological and practical applications. Also, practical suggestions are provided for teachers, educators and policy makers, followed by some limitations of the study. The next chapter presents details of the literature reviewed for addressing the research objectives of this study.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

The purpose of this research was to assess the joint influence of students' prior knowledge and classroom learning environment on their self-efficacy beliefs in achieving and achievement (see objectives in Subsection 1.5.2). To accomplish these objectives, this chapter presents a literature review in six subsections, based on a detailed online search of the Curtin University library databases and, other published books and online sources such as Google Scholar (www.scholar.google.com).

Section 2.2 covers studies involving students' mathematics achievement at the Year 10 level, which was a criterion variable of this study, and presents issues facing teachers when addressing diversity in classes and gender differences in Australia, with international comparisons.

Section 2.3 reviews studies of students' prior knowledge as a determining factor of students' achievement. Also reviewed are studies that operationalised the measurement of prior knowledge required for learning so that its influence on achievement can be investigated.

Section 2.4 reviews research related to mathematics learning environment, including constructivist theory and the philosophy of social constructivism for mathematics learning as well as studies with instruments for measuring students' perceptions of mathematics classroom learning environment, highlighting research gaps in the measurement of the construct of mathematics classroom learning environment.

Section 2.5 offers a review of studies on social cognitive theory and its application to mathematics education in the secondary schools. It explains how this theory underpins the development of the research model postulated in this study and focuses on literature related to students' mathematics self-efficacy. Also, this section

reviews studies of existing questionnaires, with their relevance for the development and validation of a new instrument for measuring the construct of self-efficacy beliefs in achieving a given set of mathematics tasks. Finally, Section 2.6 presents a chapter summary.

2.2 Students' Achievement in Mathematics

This section reviews studies of students' mathematics achievement, teaching mixed-ability groups and gender differences in achievement. Subsection 2.2.1 reviews studies dealing with students' mathematics achievement in Australia within a global perspective. Subsection 2.2.2 reviews studies of mathematics teaching and learning for ability groups (streaming *vs.* tracking). Subsection 2.2.3 reviews studies of gender differences in mathematics learning in Australia and other countries including India.

2.2.1 Students' Achievement in Mathematics at the Year 10 Level: A Global Perspective

Students' achievement is often used as a criterion variable in outcomes-focused learning environments research to improve students' learning and the quality of teaching (Aldridge & Fraser, 2008; Fraser, 2002). In mathematics education, past global research shows that students' achievement often is measured by performance on a standardised test. For example, the Programme for International Student Assessment (PISA) for 15-year-old students at the Year 10 level is administered on a triennial basis to provide national comparisons of students' achievement in mathematics according to students' background factors like gender and mathematics attitudes. Past studies in mathematics (and science) education in OECD countries cautioned about the decline in student achievement and enrolments in mathematics courses and pointed to an immediate need for curriculum changes focused on engaging all young people in mathematics. For instance, based on PISA 2012 results, Stacey and Turner (2015) concluded that, in many countries, PISA's results of mathematical literacy, when analysed for what makes mathematics education useful for most future

citizens, have been extremely influential in curriculum reviews and for improving teaching and learning.

For Australia, as well, past available research identified a decline in mathematics achievement. According to Gurria (2016), the PISA 2015 results suggest that, for the Year 10 level in Australia, the mean score in mathematics was estimated at 494 which was just above the OECD country average score of 490, but much below the leading performer (Singapore at 564). Previous PISA results suggest a steady decline in Australia's performance over the past from an estimated average of 524 points in 2003 to 514 in 2009, 504 in 2012 and 494 in 2015, which was attributed to shortcomings of classroom practices and curricula (Niemann et al., 2017; OECD, 2016a, b).

Unfortunately for India, it is difficult to make a meaningful international comparison of students' achievement because India is not in the PISA (Chhapia, 2012). However, some anecdotal evidence from the research division of the curriculum authority of India, the National Council of Education Research and Training (NCERT) provides results of India's first National Achievement Survey (NAS) for Class X in 2014-2015 (Sreekanth et al., 2015). The tests used in the NAS were based on Item Response Theory (IRT) techniques (De Ayala, 2009) and administered to a sample comprising 277,416 students in 7,216 schools across all States and Union Territories (UTs) of India. Class X results for 2015 indicated that students' mathematics mean scores were relatively low (Sreekanth et al., 2015).

The Central Board of Secondary Education (CBSE) also holds the national examination for Class X every year. According to Sreekanth et al. (2015), Class X results of mathematics achievement showed an increasing trend in the percentage of students who scored more than 40%. In summary, there is limited evidence in reviews of studies on PISA results about causes of low achievement in Australia, and available research evidence for India reveals details about neither achievement scores by ability groups for mathematics nor student-based cognitive factors that could explain the variance in achievement. The following section presents a review of studies pertaining to mathematics teaching and learning practices in mixed-ability classroom settings.

2.2.2 Mathematics Teaching and Learning: Issues of Ability Grouping

This subsection reviews studies of teaching practices in mixed-ability classrooms, such as ability grouping versus tracking adopted in Australia and a few other countries. According to Sullivan (2015, p. 243):

> …there are different forms of this stratification. The most common is when students are assessed in mathematics and then grouped according to the results on that assessment. While elsewhere different terms are used, in Australia this is termed streaming. There are also many schools which select one or more high-achieving groups, but otherwise have the rest of the groups grouped heterogeneously (commonly described as tracking). There is limited research on the extent of streaming practices in Victorian schools. In a detailed study, using an on-line survey of grouping practices in mathematics. Forgasz (2010) reports that 80 % of the 44 responding schools had some form of streaming in the years 7–10, with three quarters of the respondents indicating support for that streaming. Indeed, of the four schools which said that there was no streaming, three of the respondents were opposed to that policy. It can be inferred that the teachers who responded were overwhelmingly in favour of some streaming.

Also, in homogeneous grouping (or streaming), students might be taught different content, which could narrow the opportunities otherwise open for some students excluded in those streams; therefore, streaming could restrict equal opportunities for all. Furthermore, Sullivan (2015, p. 244) cautioned that:

> …there are significant barriers to overcoming negative effects of homogeneous grouping. In particular, students in all groups might not have the same opportunities if the curriculum is stratified and only a limited subset of the curriculum is offered to some groups, which can be exacerbated if teachers feel that skills precede other learning.

Sullivan (2015, p. 244) also argued that:

> ...tracking might result skill being emphasised to the detriment of other aspects of mathematics, such as communication, meaning and relevance. Thus, tracking encourages the placement of students in low streams, which communicates to students that their teachers think that they cannot learn. A further risk is that the 'homogeneous' grouping of students communicates to teachers that the students are indeed of like achievement, which could deter them from offering students challenging

Internationally, Oakes (1992, p. 13) reported that: "the best evidence suggests that, in most cases, tracking fails to foster the outcomes schools value", and Oakes's (2005) study concluded that low-stream classes are "deadening, non-educational environments" (p. 13). In a major meta-analysis, Hattie (2009, p. 90) concluded that "stratification, streaming, tracking, setting has minimal effect on learning outcomes and profound negative equity effects", ….and that "the quality of teaching and the nature of students' interactions are the key issues, rather than the compositional structure of the classes" (p. 90).

Forgasz (2010b, p. 66) argued that:

> …. the results are inconclusive, particularly for those at the highest levels of achievement, but for those in the middle and lower achieving mathematics classes, students may be disadvantaged with respect to achievement, and that their future mathematics and life options are likely to be curtailed.

In other words, streaming students for mathematics teaching or learning poses a threat to equity and opportunity, and is not consonant with the main goal of Melbourne Declaration of the Australian curriculum, as evident in the overarching Shape Paper that established the principles for the Australian Curriculum (ACARA, 2009, p. 5):

> All Australian governments have committed to the goals of the Melbourne Declaration, which are that Australian schooling promotes equity and excellence; and that all young Australians become successful learners, confident and creative individuals, and active and informed citizens.

An international study by Steenbergen, Makel and Olszewski (2016) involved two second-order meta-analyses and synthesised approximately 100 years of research in the United States on the effects of ability grouping and acceleration on K-12 students' academic achievement. Outcomes of 13 ability grouping meta-analyses showed that students benefited from WITHIN Class grouping (with effect sizes ranging between 0.19 SD and 0.30 SD), cross-grade subject grouping (0.26 SD) and special grouping for the gifted (0.37 SD), but did not benefit from BETWEEN Class grouping (0.04-0.06 SD); the effects did not vary for high-, medium- and low-ability students. Three meta-analyses showed that accelerated students significantly outperformed their non-accelerated same-age peers (0.70 SD) but did not differ significantly from non-accelerated older peers (0.09 SD). Three other meta-analyses that aggregated outcomes across specific forms of acceleration revealed that acceleration had a positive, moderate and statistically-significant impact on students' academic achievement (0.42 SD).

According to Law (2014), in the American education system, the problem is that tracking can lead to inequity in students' educational experiences. In Law's quantitative study, the mathematical achievement growth of students from a select elementary school (N=217) during years when students received mathematics

instruction in their regular, heterogeneously-grouped classroom was compared with that of students who received mathematics instruction in tracked classrooms. Additionally, the mathematical achievement growth of tracked students from the select school was compared with that of untracked students at a comparison school ($N = 203$). Tracking had a negative effect on students' mathematical achievement growth. Students from the select school had higher mathematical achievement growth during years when they were not tracked for mathematics instruction relative to when they were tracked for mathematics instruction. This study also revealed that students who did not undergo a change in tracking status had higher mathematical achievement growth than students whose tracking status was changed. The findings of this study suggest that the practice of tracking inhibits students' mathematical achievement growth and that schools could increase student achievement in mathematics through grouping students heterogeneously for instruction.

One way to minimise the effect of diversity on teaching or learning is either to grapple with diversity issue through streaming or grouping, or to manage diversity by reducing its impact by individualised learning plans. Askew (2015) argued that "the discourse in some schools, classrooms and policy circles frames diversity as a barrier to effective teaching or learning …. diversity must be either reduced, through practices like setting and streaming, or it must be 'managed' to reduce its impact, through practices like individualised learning experiences" (p. 129).

In summary, the teaching challenge in a mixed-ability setting lies in addressing students' individual needs without depriving any single child of opportunities that can promote self-worth and interest in the subject. As Sullivan (2015) recommends, it is preferable for all students in mixed-ability class settings to have ready access to challenging learning opportunities in the interest of equity. Streaming or tracking students has its own limitations and deprives some students of suitable learning opportunities, as reported by many international studies (Oakes, 1992; Hattie, 2009; Steenbergen et al., 2016). As Askew (2015) recommends, diversity can be reduced by managing differences between learners, such as by promoting teaching practices that encourage individualised learning experiences in a mixed-ability classroom setting.

2.2.3 Gender Differences in Mathematics Achievement

Regarding the gender gap in mathematics achievement, Baker and Jones (1993, p. 91) clarified that "we use the term gender to refer to social roles of the two sexes and the term sex difference to mean a difference in behaviour between the sexes. This is emerging as an accepted practice (Giele 1988)". They examined the question: "Is variation in gender stratification of opportunity related to sex differences in the performance of academic mathematics?" (p. 93) and found that "there is cross-national variation in the performance of mathematics and that it is related to variation in the gender stratification of educational and occupational opportunities in adulthood, that sex differences have declined over time, and that school and family factors leading to higher mathematical performance are less stratified by gender when women have more equal access to jobs and higher education" (p. 91). Baker and Jones (1993) argued that female students, who are faced with less opportunities, might see mathematics as less important for their future because of what they are told by teachers, parents and friends.

Spencer, Steele and Quinn (1999, p. 4) noted that "when women perform math, unlike men, they risk being judged by the negative stereotype that women have weaker math ability. We call this predicament stereotype threat and hypothesize that the apprehension it causes may disrupt women's math performance".

Hill et al. (2010, p. 38) observed that "Two stereotypes are prevalent: girls are not as good as boys in math, and scientific work is better suited to boys and men. Negative stereotypes about girls' and women's abilities in mathematics and science persist despite girls' and women's considerable gains in participation and performance in these areas during the last few decades."

To investigate gender differences in mathematics achievement, Else-Quest, Hyde and Linn (2010) meta-analysed two major international data sets, the 2003 Trends in International Mathematics and Science Study and the Programme for International Student Assessment, representing 493,495 students of 14-16 years of age, to estimate the magnitude of gender differences in mathematics achievement, attitudes,

and affect across 69 nations throughout the world. After analysing PISA data, Else-Quest et al. (2010, p. 103) reported that:

> …consistent with the gender similarities hypothesis, all of the mean effect sizes in mathematics achievement were very small (d < 0.15); however, national effect sizes showed considerable variability (d = -0.42 to 0.40). For TIMSS mathematics, the mean weighted effect sizes (d) were estimated also as well as the content domains of Algebra, Data, Geometry, Measurement, and Number. For example, for mathematics achievement, the weighted mean effect size for the gender difference in performance was d=0.01, indicating that boys and girls performed similarly overall.

For Australia, Else-Quest et al. (2010) indicated negligible gender difference in mathematics achievement at the Year 10 level. Based on participation of 6,171 girls and 6,380 boys, they reported that "the effect size was found to be 0.06 SD, which indicates a negligible difference in performance between boys and girls" (p. 118).

More recently, in an online survey in Australia, Forgasz and Leder (2015, p. 103) sought public views regarding prevalent stereotypes of mathematics and English as a male-female domain and reported that "…many in this group do not gender stereotype mathematics or English. Among those who do, however, the traditional stereotypes persist. Thus, challenges remain to counter such stereotyping and enhance the perception, and reality, of mathematics as an inclusive field appropriate for all".

Regarding students' mathematics achievement in India at a comparable level, India was not included in the PISA (Chhapia, 2012). However, other research sources, namely, the National Curriculum Framework (NCF-2005) as cited by the curriculum authority in India, provide anecdotal evidence (NCERT, 2012, 2015).

India is a multicultural developing country with numerous diversity issues of inclusive education in terms of social, cultural, linguistic, religious and economic aspects. The national literacy rate nevertheless increased from 64.83% in 2005 to 74.04% by 2015, although elementary education is a fundamental right for every child as guaranteed by the Indian Constitution (Government of India, 2005, 2015). According to the Census (Government of India, 2015), there exist widespread differences among states / Union Territories (UTs), with the literacy rate being the lowest in the state of Bihar (63.8%) and the highest in the state of Kerala (93.91%).

Female literacy rates were lower than male literacy rates in general, especially in Rajasthan.

The Indian curriculum addresses the persistence of stereotypes in terms of gender, caste and physical and intellectual disability. NCF-2005 (NCERT, 2012) spells out strategic directions for school administrators, teachers and other stakeholders for managing such differences. About managing the impact of diversity, the NCF-2005 (NCERT, 2012) specifies modalities:

> It is important to create an inclusive environment in the classroom for all students, especially those who are at risk of marginalisation, for instance, students with disabilities…… Differences between students must be viewed as resources for supporting learning rather than as a problem. Inclusion in education is one of the components of inclusion in society. (p. 16)

However, the NCERT study conducted by Sreekanth et al. (2015) for the Class X (or Year 10 level) on an all-India basis provides evidence for nonsignificant gender differences in mathematics achievement.

In summary, the study of gender differences in mathematics has received researchers' continued attention for different reasons. While recent research evidence in Australia (Forgasz & Leder, 2015) pointed to the persistence of traditional stereotypes of mathematics being perceived as a male domain, the Else-Quest et al.'s (2010) meta-analytical study identified that the gender differences in mathematics achievement were a cause of underrepresentation of women in careers in science, technology, mathematics and engineering. Using 2003 PISA data, Else-Quest et al. (2010) reported effect sizes for gender differences in mathematics at different stages of learning, and observed that national effect sizes were more varied from -0.42 SD to 0.40 SD, with the mean effect size in mathematics achievement across 69 nations, however, being small (0.15 SD) and for Australia being smaller at 0.06 SD. However, effect sizes were estimated neither in the studies reviewed for gender differences for a topic within the mathematics content domain at the year 10 level nor by ability groups, which are research gaps addressed in my research study.

The following section presents a review of studies of the construct of prior knowledge required for learning and its role in influencing students' achievement.

2.3 Students' Prior Knowledge in Mathematics

This review of past research about students' prior knowledge as an important determinant of achievement is organised into three subsections: prior knowledge as a construct in Subsection 2.3.1; operationalisation of prior knowledge in Subsection 2.3.2; and assessment of prior knowledge in Subsection 2.3.3.

2.3.1 Students' Prior Knowledge as a Construct

As students progress to higher levels in schooling, their prior knowledge required for learning also builds up and assumes greater importance for learning a new topic. According to Dochy (1992), prior knowledge is conceptualised as the learner's content knowledge related to a domain studied, which is present before the implementation of instruction. In his assimilation theory, Ausubel (1968) described the learner-controlled process as assimilation and argued that meaningful learning occurs when learners connect new information to their existing knowledge structure. According to this theory, assimilation has three components: i) learners' prior knowledge at its core, ii) meaningful material and iii) learners' intent and ability to use their prior knowledge.

Schneider, Korkel and Weinert (1989) showed that prior knowledge can compensate for low aptitude, while high aptitude could not compensate for low prior knowledge. Individual differences in prior knowledge are a primary determinant of achievement (Dochy, 1996; Tobias, 1994). Therefore, students who lack relevant prior knowledge face difficulties in learning new information (Ausubel, 2000).

2.3.2 A Framework for Operationalising Prior Knowledge

For operationalising prior knowledge, many cognitivists recognise conditional knowledge and two types of declarative knowledge and procedural knowledge (Alexander, Schallert, & Hare, 1991; Paris, Lipson, & Wixson, 1983). Alexander et al. 1991, p. 323) clarified:

> Whether we are speaking of content knowledge, linguistic knowledge, or any other form of knowledge, we hold to the premise that any of these forms can contain declarative, procedural, or conditional knowledge. When we know something (be it content, linguistic, or otherwise), we can know not only information about it (declarative knowledge) but also how to use such knowledge in certain processes or routines (procedural knowledge). We can also understand when and where this knowledge would be applicable (conditional knowledge). It is important to remember that these three types of knowledge are distinct; the acquisition of knowledge in one form does not automatically and immediately guarantee knowledge in the other forms. Thus, it is certainly possible to know the what of a thing without knowing the how or when of it.

Dochy (1992) defined declarative knowledge as the accumulation of facts and concepts that come to the surface by recognition or reproduction. Anderson (1995) refers to declarative knowledge as 'knowing that', whereas procedural knowledge as 'knowing how', which can be identified in assessment through production or application. Following this, Hailikari, Nevgi and Lindblom-Ylänne (2007) provided a framework for operationalising prior knowledge in mathematics learning that has four types of prior knowledge as shown below in Table 2.1. This framework offers a rationale for identifying tasks that could represent the four components of students' prior knowledge required for learning a mathematics topic.

2.3.3 Assessment of Prior Knowledge

Many researchers have investigated plausible relationships between students' prior knowledge and learning outcomes (Dochy, 1994). For example, in two long-term longitudinal studies of children aged 4-12 years, Weinert and Helmke (1998) showed that correlations between prior knowledge and mathematics performance remained strong, even when intelligence was partialled out. The influence of prior knowledge on performance improvement was strong, while the influence of intelligence was weak as a source of difference in performance.

Hallinan and Sorenson (1987) considered knowledge and sex when organising students into ability groups for instruction. One of the important findings was that students' prior knowledge, as measured by a grade-equivalent mathematics score at the beginning of the year, had a strong positive effect on group assignment. That is, "the higher a student's standardised test score in mathematics, the higher the likelihood that the student will be assigned to the high-ability group" (Hallinan & Sorenson, 1987, p. 66). This finding suggests that students' prior knowledge can be used as a reasonable

basis for organising ability groups within a mixed-ability classroom setting. The cognitive needs of students in each ability group can be addressed by teachers providing them with suitable learning opportunities and tasks (Askew, 2015; Sullivan et al., 2014).

Hailikari, Nevgi and Komulainen (2007) examined relationships between prior knowledge, academic self-beliefs and previous study success in predicting the achievement of 139 students in a university mathematics course using structural equation modelling. Hailikari et al. (2007, p. 59) reported that "domain-specific prior knowledge was the strongest predictor of student achievement", while "prior knowledge and previous study success showed the strongest positive correlation with the final grade" (p. 67). Academic self-beliefs correlated with previous study success, had a strong direct influence on prior knowledge test performance, and predicted student achievement only indirectly via prior knowledge.

Table 2.1: A Framework for Operationalising Students' Prior Knowledge in Mathematics

Knowledge Components		Operationalisation
Declarative Knowledge		
1)	Knowledge of facts	Free recall; enumerating concepts related to the subject matter
2)	Knowledge of meanings, concepts	Open questions giving definitions to the recalled concepts
Procedural Knowledge		
1)	Integration of knowledge	Questions about interrelations between different mathematical concepts
2)	Application of knowledge	Mathematical problem-solving tasks

Note: Adapted from Hailikari, Nevgi and Lindblom-Ylanne (2007, p. 325).

Hailikari, Nevgi and Lindblom-Ylänne (2007, p. 326) measured "prior knowledge with six mathematical problem-solving tasks that concentrated on optimal-requisite prior knowledge before the start of the course. These problem-solving tasks required the ability to apply knowledge to solve mathematical tasks".

Hailikari, Nevgi and Lindblom-Ylänne (2007, p. 323) identified that "almost all educational studies have acknowledged the significant and positive influence of prior knowledge on learning (Ausubel, 1968, 2000; Bloom, 1976; Dochy, 1992; Dochy et al., 2002; Thompson & Zamboanga, 2004). Knowledge that the learner already possesses about a particular subject influences the acquisition of new knowledge and

all different phases of information processing (Dochy et al., 2002)". Hailikari et al. 2007) claim that their findings were consistent with previous research and indicated that "the level of prior knowledge is significantly related to student achievement" (p. 330).

Hailikari et al. (2008, p. 1) examined "how different types of prior knowledge (declarative and procedural) impact student achievement and how prior-knowledge assessment can be used as an instructional design tool". By using the knowledge framework shown in Table 2.1, Hailikari, Nevgi and Lindblom-Yla¨nne (2007) measured the components of prior knowledge (declarative and procedural). Instructors and students had mainly positive reactions towards the prior-knowledge tests. The authors reported that "Prior knowledge from previous courses significantly influenced student achievement. Procedural knowledge was especially related to student achievement" (p. 1). Learning environment research is reviewed in the following section.

2.4 Classroom Learning Environment Research

This section reviews studies of learning environment research in two subsections, starting with theories of learning including constructivist theory, the philosophy of social constructivism and socio-cultural perspectives of mathematics learning in Subsection 2.4.1. Then, a review of studies about relevant instruments for measuring students' perceptions of learning environment in mathematics classrooms is the focus of Subsection 2.4.2.

2.4.1 Theories of Mathematics Learning

Zone of Proximal Development (ZPD) offers an important perspective of how and when learning occurs. According to Vygotsky (1978, p. 86), ZPD is "the distance between a child's actual developmental level as determined by independent problem solving and their higher level of potential development as determined through problem

solving under adult guidance or in collaboration with more capable peers". Walshaw (2016) notes that learning is most effective when the student is introduced to new concepts that are on the cusp of emergence for the student. The ZPD was not conceptualised as a permanent state but, rather, as a stage towards independent knowing or acting.

A concept often used in association with the ZPD is *scaffolding*, a process in which, adult assistance enables a student to solve a problem for carrying out a task or to achieve a goal beyond his/her unassisted efforts at that point in time (Wood, Bruner, & Ross, 1976).

Interrelated with scaffolding is constructivist theory, which suggests that individual learners construct mental models to understand the world around them. According to Moll et al. (1980, p. 54), "Vygotsky's (1978) idea is that learning occurs in a zone of proximal development". That is, in a ZPD, a learner (e.g. student) requires an adult's (teacher or a peer) help to construct learning. Thus, learning is a situated social activity that operates only when the child is interacting with people in his/her environment and in co-operation with his/her peers.

Harlow, Cummings and Aberasturi (2006, pp. 45-47) discussed constructivism by examining the learning theories of Jean Piaget and Karl Popper, whose thesis was that "the construction of new knowledge developed from a state of disequilibrium, in which the learner was confronted with an object which could not be assimilated into prior knowledge and thus reconstruction occurred".

Brooks (1990, p. 69) elaborated that "The intellectual opportunities the teacher offers the students are carefully constructed invitations that maximize the *possibility* that new conceptual learning will occur".

According to von Glasersfeld (2000), the epistemological belief of a radical constructivist is "one that holds that knowledge is under all circumstances constructed by individual thinkers as an adaptation to their subjective experience" (p. 4). It thus connects students' prior knowledge to learning a new mathematics topic in the classroom. The constructivist teacher presents new information in the context of a problem that introduces disequilibrium, but that students find interesting and can solve

with some effort and support. This process is experienced by students across all grade levels and ability groups.

For instance, Broza and Kolikant (2015) considered 11 fifth-grade low-ability students and monitored their progress when teachers adopted a scaffolding strategy to teach decimal subtraction. They found that nine students showed significant improvement, but that their learning processes were inconsistent and difficult to predict. This study identified what characterises meaningful learning of mathematics among low-achieving students (LAS) and highlighted challenges posed for contingent teaching, the adaptive core of scaffolding. The setting was an extracurricular program for teaching meaningful mathematics to LAS through a combination of learning in context, interactive computerised activities and contingent teaching. Micro-genetic analysis involved lesson transcripts, videotaped computer activities and worksheets, as well as data from teachers' and students' pre- and post-program interviews.

Ernest (1991, 1998) advanced his philosophy of social constructivism by arguing that mathematics is a language that is socially constructed by learners. Following Ernest's philosophy, Wood and Turner-Vorbeck (2001) offered a socio-cultural perspective for analysing the roles that teachers and students play as both explainers and listeners, sharing their responsibility for thinking and learning by social interactions in the classroom.

Albert (2012) provided a social perspective of constructivism in which mathematics learning is developed in stages by teachers or significant 'others' by asking intriguing questions while facilitating peer interaction. Albert (2012) elaborated Vygotsky's sociocultural dimension in the practice of scaffolded instruction, as an example, for assisting teachers in developing their understanding of pedagogical content knowledge in mathematics.

2.4.2 Theoretical Rationale for Measuring Classroom Learning Environment

Past studies recognise the importance of constructivist theory for developing instruments to measure students' perceptions of classroom learning environment in different countries (Aldridge, Fraser, Taylor, & Chen, 2000; Taylor, Fraser, & Fisher,

1997). To measure students' perceptions of the extent to which constructivist approaches are present in classrooms, Taylor et al. (1995) and Taylor et al. (1997) developed the Constructivist Learning Environment Survey (CLES). The original version of the CLES (Taylor & Fraser, 1991) was based largely on a psychosocial view of constructivist reform that focused on students as co-constructors of knowledge. As observed by Fraser (1998), the framework supporting the CLES was found to be weak. Aldridge et al. (2000) developed a new version of the CLES for the science domain from the perspective of critical constructivism (Taylor, 1996), after considering the socio-cultural constraints on the cognitive constructive activity of the individual learner such as: the degree of personal relevance in their studies; whether students have shared control over their learning; the degree to which students feel free to express concerns about their learning; the degree to which students are able to interact with each other to improve their understanding; and the extent to which science is viewed as ever changing (Aldridge et al., 2000, pp. 38-39).

Aldridge et al. (2000, p. 39) designed the CLES with five key dimensions of critical constructivist learning environment: " Personal Relevance (extent to which teachers relate science to students' out-of-school experiences); Student Negotiation (extent to which opportunities exist for students to explain and justify to other students their newly-developing ideas and to listen and reflect on the viability of other students' ideas); Shared Control (extent to which students are invited to share with the teacher control of the learning environment, including the articulation of their own learning goals, the design and management of their learning activities and the determination and application of assessment criteria); Critical Voice (extent to which a social climate has been established in which students feel that it is legitimate and beneficial to question the teacher' s pedagogical plans and methods and to express concerns about any impediments to their learning); Uncertainty (the extent to which opportunities are provided for students to experience scientific knowledge as arising from theory-dependent inquiry, involving human experience and values, evolving and non-foundational, and culturally and socially determined)".

Aldridge et al. (2000) used the 30-item CLES with a five-point response scale of Almost Always, Often, Sometimes, Seldom and Almost Never, after omitting

negative items and rearranging items within a specified scale in blocks rather than randomly. They validated CLES by administering to 1,081 students from 50 classes in Australia and 1,879 students from 50 classes in Taiwan.

Aldridge et al.'s (2012) study incorporated these dimensions along with others in developing the Constructivist-Oriented Learning Environment Survey (COLES) that has 88 items spread across 11 dimensions that include Personal Relevance, Shared Control and Uncertainty.

In Victorian schooling, all teachers must adhere to Six Principles of Learning and Teaching (State Government of Victoria, 2012) which, when compared with COLES dimensions as shown in Table 2.2, reinforce the relevance of 11 key dimensions. Aldridge et al.'s (2012) study validated the COLES by administering it to large sample of 2043 Years 11-12 students from 147 classes of nine schools in Western Australia.

Past research (Germann, 1994; Henderson, Fisher, & Fraser, 1995; Rawnsley, 1997) emphasised that learning environments typically are positively associated with attitudes and achievement. For example, Aldridge et al. (2012) reported associations between classroom learning environment dimensions and learning outcomes, including students' attitudes and achievement.

Extensive literature reviews of classroom environment research contributed by Fraser (2002, 2012, 2014) provide compelling evidence that the What Is Happening In this Class? (WIHIC) questionnaire has been used successfully and adopted in different curriculum domains in different countries.

For Indian classrooms, Koul (2003, 2009) provides some evidence about science learning environments by administering the Questionnaire on Teacher Interaction (QTI), the What is Happening In this Class? (WIHIC) and an attitude scale to 1,021 Years 9-10 students from 32 science classes in seven different co-educational private schools in one of the Indian states, Jammu and Kashmir.

The data were analysed to support the reliability and validity of each scale. The simple correlations between students' perceptions of learning environment, their attitudes and cognitive achievements ranged from 0.17 to 0.38.

Table 2.2: Mapping Six Principles of Learning and Teaching to COLES and Moos' Dimensions

Principles of Learning and Teaching P-12*	Dimensions in COLES**
1. The learning environment is supportive and productive.	Student Cohesiveness Teacher Support
2. The learning environment promotes independence.	Teacher Support Involvement Cooperation Equity
3. Students' needs, backgrounds, perspectives and interests. are reflected in the learning.	Differentiation Personal Relevance
4. Students are challenged and supported to develop deep levels of thinking and application.	Involvement Task Orientation
5. Assessment practices are an integral part of teaching and learning.	Formative Assessment Clarity of Assessment Criteria & Feedback
6. Learning connects strongly with communities and practice beyond the classroom.	Personal Relevance

Notes: * State Government of Victoria (2012) and ** Aldridge et al. (2012)

Similarly, Gupta and Fisher (2012) conducted a study in the science domain among students from the same state of Jammu and Kashmir in India by using a modified form of Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) that had been pre-validated by Aldridge et al. (2008). Analysis of data from 705 students from 15 classes provided evidence for the reliability and validity of the questionnaire in Indian science classroom settings, as well as information about gender differences and associations between students' perceptions of their technology-supported learning environments and three learner outcomes (attitude towards science, academic efficacy and academic achievement).

These studies pioneered the field of classroom learning environment research in India, but they were confined to the science domain. However, the COLES could be extended to mathematics classrooms in India to fill a research gap. The following section reviews studies of social cognitive theory and self-efficacy theory to identify a

suitable instrument to measure self-efficacy beliefs in achieving mathematics tasks, which was a major research objective of my research study.

2.5 Social Cognitive Theory

In relation to research questions 4 and 5 of Subsection 1.5.1, relevant studies about social cognitive theory and self-efficacy are reviewed in the following five subsections. Subsection 2.5.1 reviews studies that offer a theoretical rationale for the relationships within the framework of the research model shown in Figure 1.1. Subsection 2.5.2 reviews studies about the relationship between self-efficacy, prior knowledge, classroom learning environment and achievement. Subsection 2.5.3 reviews studies of self-efficacy at a task-specific level, while Subsection 2.5.4 deals with past research about the mediational role of self-efficacy in influencing achievement. Finally, Subsection 2.5.5 reviews studies that employed self-efficacy instruments as a classroom tool for measuring self-efficacy beliefs and the calibration of self-efficacy judgements into efficacy-expectancies.

2.5.1 Theoretical Framework for the Present Research

Bandura's (1986, p. 22) social cognitive theory construes "human functioning as a series of reciprocal interactions between personal influences, environmental features and behaviours". As shown schematically in Figure 2.1, a hallmark of social cognitive theory is that human functioning involves reciprocal interactions between behaviours, environmental events, cognition and other personal factors. It describes human learning in terms of the triadic relationship between personal, environmental and behavioural determinants. The underpinning of social cognitive theory is human agency (i.e. individuals are agents who are proactively engaged in their own development and can make things happen by their actions). Bandura (1986a, p. 25) theorised that "central to human agency is the idea that individuals possess self-beliefs that enable them to exercise a measure of control over their thoughts, feelings and actions, and that, what people think, believe, and feel affects how they behave".

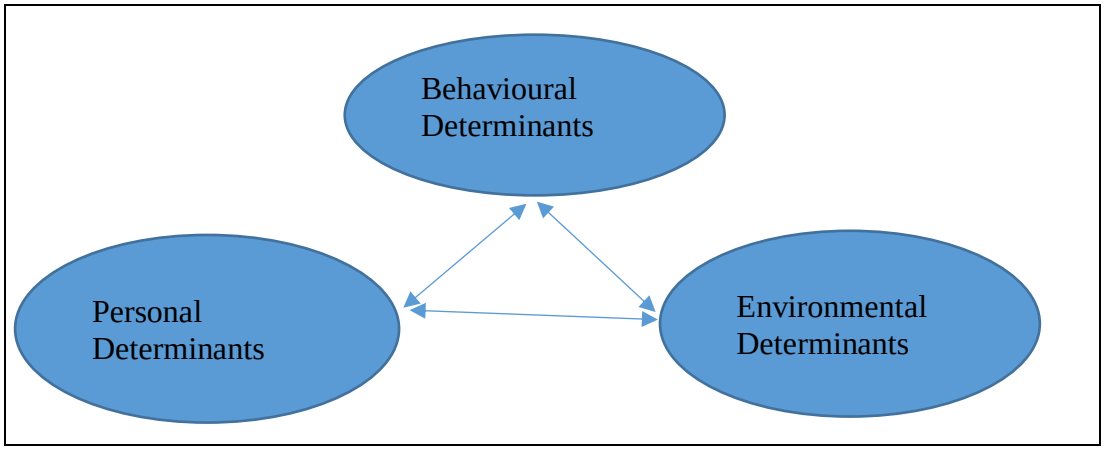


Figure 2.1: Triadic Interrelationships between Behavioural, Personal and Environmental Determinants of Human Functioning (adapted from Bandura, 1986)

In the field of education, Urdan and Schoenfelder (2006, pp. 340-341) argue that, "schools and classrooms are, by definition, social environments. Within any single classroom, students have social interactions and build social relationships with their teacher, with close friends, and with their non-friend classmates. … The relationships between students and teachers influences classroom climate, which we define as the general class atmosphere including attitudes towards learning, norms of social interaction, acceptance of ideas and mistakes, and learning structures set by the teacher. Teachers are responsible for regulating the academic environment, including the material covered, approaches to learning presented, and the manner in which individuals communicate within the classroom…. Teachers can alter school and classroom structures that underpin students' success (environmental factors), work to improve students' emotional states, their beliefs about themselves and their habits of personal, environmental and behavioural determinants (personal factors). Such interventions can improve their academic skills and strategies for learning behaviour".

After reviewing Bandura's (1986) seminal work, Locke (1987, p. 170) commented that:

> ...self-efficacy is shown to be tied to performance effectiveness in a wide range of situations including: dealing with feared situations (phobias), coping with stress, sports performance, breaking habits like smoking, recovery from heart attacks, and motivation to set and attain high goals.

The following subsection reviews studies about the relationships between self-efficacy and i) prior knowledge, ii) classroom learning environment and iii) achievement.

2.5.2 Relationships between Self-Efficacy, Prior Knowledge, Classroom Learning Environment and Achievement

Schunk and Pajares (2004, p. 115) observed that, "in educational settings, self-efficacy is not an isolated construct but rather an integral component of social-cognitive theory. Students' social environment can influence both attitudes and behaviours. The interaction between self-efficacy and environment is evident in classrooms". For example, the types of questions that teachers ask students, how students are grouped for instructions, and the feedback that teachers give to students about their performance are environmental variables that can affect students' self-efficacy.

Additionally, Urdan and Schoenfelder (2006, p. 341) argued that "teachers, as an integral component of the classroom environment, can inspire students by creating a favourable classroom environment where students feel more personally efficacious and motivated and, therefore, are willing to work harder to succeed".

Various researchers investigated the role of self-efficacy beliefs in students' mathematics classroom learning environment and achievement. Collins (1982), as cited by Schwarzer (2014), examined self-efficacy beliefs as a causal factor in human functioning. To test this proposition, the author selected children who judged themselves to be of high or low mathematical efficacy at each of three levels of mathematical ability. They were then given difficult problems to solve. At each level of mathematical ability, students who regarded themselves as efficacious were quicker to discard faulty strategies, solved more problems, chose to rework more of the problems that they failed, and did so more accurately than those of equal ability who doubted their efficacy. Perceived self-efficacy thus exerted a substantial independent effect on performance. Positive attitude toward mathematics was better predicted by perceived self-efficacy than by actual ability.

Hackett (1985) reported results of a path analysis of mathematics scores on American College Test (ACT), indicating that mathematics self-efficacy contributed more than sex, years of high school mathematics or mathematics anxiety in predicting the choice of a mathematics-related college major.

In a comparative study of classrooms in Australian and British secondary schools, Dorman and Adams (2004) examined the relationship between academic self-efficacy beliefs and classroom learning environment. They used the 70-item WIHIC, with its seven dimensions, and three scales from CLES. They used a seven-item scale developed by Midgley and Urdan (1995) and Roeser et al. (1996), with one of the items being reversed, to assess students perceived academic competence at mathematics-related tasks. The response format for all academic efficacy items was a 5-point scale from 1 (not at all true) to 5 (very true). For validation, they administered these instruments to large sample of 2631 students.

Dorman and Adams' (2004) study reported satisfactory validity for all scales and positive and statistically-significant correlations between academic self-efficacy and each of the ten classrooms learning environment dimensions. Multiple regression of academic self-efficacy on all ten dimensions indicated a R^2 value of 19.4% of the total variance in self-efficacy. However, only six out of ten regression coefficients were positive and significant, with their estimates ranging from as low as 0.06 for Equity to as high as 0.27 for Task Orientation. The self-efficacy effects were statistically nonsignificant for the other dimensions of Student Cohesiveness, Teacher Support, Personal Relevance and Shared Control.

In a California-based study, Fast, Lewis and Bryant (2010) assessed the effect of the perceived classroom environment on mathematics self-efficacy, as well as the effect of mathematics self-efficacy on standardised mathematics test performance, among the upper-elementary school students (N=1,163). A series of two-level structural models estimated by the authors revealed that students who perceived their classroom environments as more caring, challenging and mastery-oriented had significantly higher levels of mathematics self-efficacy, and that higher levels of mathematics self-efficacy positively predicted mathematics performance.

Hailikari, Nevgi and Komulainen's (2007) study aimed to explore the relationships between previous study success, domain-specific prior knowledge, and academic self-beliefs in predicting achievement among 139 university mathematics students. Structural equation modelling was used for predicting student achievement.

Hailikari, Nevgi and Komulainen (2007, p. 59) defined academic self-beliefs as, "an individual's beliefs about his or her attributes and abilities as a learner", and they measured academic self-beliefs using students' responses to survey statements based on Valentine, DuBois and Cooper (2004) and Niemi et al. (2003). Hailikari et al. (2007) assessed students' confidence in achieving in terms of expectation of success and students' self-reports of mathematics ability. Their responses were rated on a five-point Likert scale from 1 (Completely disagree) to 5 (Completely agree).

Hailikari et al. (2007) argue that students' academic self-efficacy beliefs and prior knowledge together influence achievement. Academic self-beliefs strongly correlated with previous study success and had a strong direct influence on prior knowledge test performance. However, self-beliefs predicted student achievement only indirectly via prior knowledge. The results imply that both prior knowledge and self-beliefs should be considered when considering instructional support issues, because they can provide valuable insights about the future performance.

But, the above studies did not consider self-efficacy at a task-specific level, which was recommended by Bandura (1980) as described below.

2.5.3 Measurement of Task-Specific Self-Efficacy Beliefs

Bandura (1980) made a strong case that self-efficacy should be investigated for individual tasks rather than generally for aggregate performance on different tasks, because of its behavioural and situational specificity. Bandura (1986b) emphasised that self-efficacy is more predictive of future performance on a task than such global indicators as confidence in learning (Fennema & Sherman, 1976). Norwich (1986) provided a framework for assessing perceived mathematics self-efficacy beliefs.

Following Bandura (1986b), Hackett and Betz (1989) defined perceived mathematics self-efficacy as "a situational or problem-specific assessment of an

individual's confidence in her or his ability to successfully perform or accomplish a [mathematics] task or problem" (p. 262). Hackett and Betz (1989) examined the role of students' self-efficacy as a predictor of performance on related tasks in a mathematics test. Based on their previous studies (Hackett, 1985; Hackett & Betz, 1981), the major hypotheses of Hackett and Betz's (1989, p. 263) study were:

- The mathematics self-efficacy/mathematics performance correspondence will be stronger for men than for women. That is, women's self-efficacy expectations will be unrealistically low compared to men's.

- Self-efficacy about specific mathematics problems will be related to actual performance on an equivalent set of problems. Theoretical predictions about the self-efficacy/performance correspondence have been confirmed at a global level, that is, general mathematics self-efficacy has been found to be strongly related to the ACT mathematics achievement scores (Hackett, 1985). However, this self-efficacy/performance relationship has not been evaluated at a more task-specific level.

Hackett and Betz (1989) investigated the relationship between mathematical performance, mathematics self-efficacy, attitudes towards mathematics and choice of mathematics majors by administering five pre-validated instruments to 153 women and 109 men from undergraduate courses:

a) Background and Career-Plans Questionnaire
b) Mathematics Self-Efficacy Scale (Betz & Hackett, 1983) comprising three subscales: (i) the Mathematics Tasks subscale, involving 'everyday' mathematics tasks (e.g., balancing a chequebook) (ii) the Mathematics Courses subscale, consisting of 16 mathematics-related college courses and (iii) the Mathematics Problems subscale consisting of 18 arithmetic, algebra and geometry problems
c) Mathematics Performance Scale (Dowling, 1978)
d) Mathematics Attitude Scales (Fennema & Sherman, 1976)
e) Bem's (1974) Sex-Role Inventory for gender differences.

Using students' responses to various items of the above scales, Hackett and Betz's (1989) study revealed that mathematics performance was correlated with mathematics self-efficacy moderately strongly and statistically significantly, ranging from 0.36 to 0.49, while no support was found for Hackett and Betz's (1981)

hypothesis that women's mathematics self-efficacy expectations are unrealistically low compared to men's.

The study by Pajares and Miller (1995) also concluded that task-specific self-efficacy beliefs are the most important determinants of students' mathematics achievement, and that they mediate previous study success.

2.5.4 Mediational Role of Self-efficacy

Many studies identified the mediational role played by students' self-efficacy in achievement and decisions about mathematics-related careers (Pajares & Urdan, 2006; Ranadhawa et al. 1993; Schunk, 1995; Siegel, Galassi, & Ware, 1985). For example, Siegel, Galassi and Ware (1985) compared alternative explanatory models of mathematics performance and reported that the self-efficacy model was superior to a mathematics aptitude / mathematics anxiety model in predicting college students' performance scores on the scholastic achievement test (SAT). Mathematics self-efficacy accounted for a much higher percentage of the variance in mathematics scores than the combined effects of mathematics anxiety and sex. Siegel, Galassi and Ware's (1985) study suggests that mathematics self-efficacy is significantly correlated both with i) attitudes towards mathematics and ii) the extent to which college students select mathematics-related college majors. Furthermore, the mathematics self-efficacy of men was significantly stronger than that of women.

Randhawa, Beamer and Lundberg (1993) also examined the mediation role of self-efficacy beliefs using a structural model of mathematics achievement. They administered two attitudes and three mathematics self-efficacy scales to 117 males and 108 female high-school seniors. Teacher-assigned scores for a selected mathematics course were also obtained. The covariance matrices of boys and girls were analysed with a 2-group LISREL procedure. The model specified mathematics self-efficacy as a mediator between mathematics attitude and achievement, and it provided a good fit to the data for both boys and girls.

The following section reviews studies that used self-efficacy instruments as classroom tools for assessing efficacy expectancies and predicting students' achievement.

2.5.5 Self-Efficacy Instruments as Classroom Tools for Assessing Efficacy-Expectancies and Predicting Achievement

Some studies involved designing instruments to measure students' perceptions of self-efficacy beliefs at different levels and mainly for addressing three types of issues (Pajares & Kranzler, 1995; Pajares & Miller, 1995):

a) To measure self-efficacy associations with learning outcomes including mathematics attitudes, anxiety, aptitude for learning and mathematics achievement.
b) To investigate the mediational role of self-efficacy beliefs in quantifying the influence of cognitive factor such as students' prior knowledge on achievement.
c) To calibrate students' self-efficacy judgements into efficacy- expectancies that could be used for predicting achievement and identifying underachievers or students with lower efficacy- expectancies.

Pajares and Kranzler (1995) observed that a student's self-assessment of self-efficacy beliefs can be used as an explanatory variable in predicting performance and thus, suggested a procedure for calibrating how well his/her self-efficacy judgements relate to performance on accompanying tasks. In doing so, it is pertinent and useful to distinguish between *self-efficacy expectancy* and *outcome expectancy*. Bandura's (1977, p. 193) self-efficacy theory maintains that a self-efficacy expectancy is the belief that the person is or is not capable of performing the requisite behaviour:

> …the expectancy that performing a response will produce a desired outcome (response-outcome or "outcome" expectancy) as against the individual's expectancy that he can perform that response (self-response or "self-efficacy" expectancy.

There have been numerous studies supporting self-efficacy expectancies (Bandura, Reese & Adams, 1982; Davis & Yates, 1982; Maddux & Rogers, 1983). Bandura (1977, p. 194) asserted that further calibration capabilities are equally important as efficacy-expectancy:

> … how perceived self-efficacy influences performance is not meant to imply that expectation is the sole determinant of behaviour. Expectation alone will not produce desired performance if the component capabilities are lacking. Moreover, there are many things that people can do with certainty of success that they do not perform because they have no incentives to do so. Given appropriate skills and adequate incentives, however, efficacy expectations are a major determinant of people's choice of activities, how much effort they will expend, and of how long they will sustain effort in dealing with stressful situations.

Furthermore, for gauging the congruence between self-efficacy judgments and actions, Bandura (1977) recommended a micro-analytical strategy for studying calibration which requires a hierarchy of tasks developed.

However, Hackett and Betz (1989, p. 266) claim for their study: "Because of the differences in the tasks involved, it was not possible to analyse the data gathered in this study precisely as Bandura has analysed the self-efficacy/action relationship in past research on animal phobias; however, it was possible to conduct a microanalysis by sex of the mathematics self-efficacy and mathematics performance relationship using a somewhat different procedure".

Thus, Hackett and Betz (1989) conducted a microanalysis of self-efficacy vs. performance correspondence. The method of estimation used for the micro analysis as well as the findings that emerged are important for informing teaching practice. Their method was based on previous research by Dowling (1978) and involved computing deviation scores (D scores) between mathematics self-efficacy and performance by separately transforming raw scores on the scales to standardised scores (*z*-scores), and then subtracting the resulting *z*-scores of performances from the *z*- scores for efficacy expectancy for comparable items. Mean D scores, calculated for each student by gender were classified into three types as shown in Table 2.3. Mean D scores less than -2 were ignored.

Table 2.3: Mean Score of Deviations of Self-Efficacy Expectancies from Performance

Mean D-score	Degree of Confidence
-2.00 - 0.00	Under-confidence in performance
0.00 - 2.00	Congruence between self-efficacy and performance
+2.00 -	Overconfidence in performance.

The mean D score, as an index of confidence, represents an average difference between students' self-efficacy in achieving specified tasks and performance in

solving those tasks. Hackett and Betz's (1989) study found that overall, only 35% of students had their D scores in the congruent range, 48% in the overconfident range and only 18% were in the underconfident range. For women, 43% were classified as overconfident and 18% as underconfident but, for men, 54% were classified as overconfident and 16% as underconfident.

However, a chi-square analysis of the data did not indicate a significant relationship for females, and thus did not support Hackett and Betz's (1981) proposition that women's mathematics self-efficacy expectations are unrealistically low when compared to their performance, whereas the mathematics self-efficacy / mathematics performance analyses support the Bandura's (1977) hypothesis of a positive relationship between self-efficacy and performance.Furthermore, Hackett and Betz (1989) reported that both mathematics performance and mathematics self-efficacy were significantly and positively correlated with attitudes towards mathematics, sex-role orientation and mathematics-related major. Regression analyses supported the superiority of mathematics self-efficacy over mathematics performance and achievement variables in predicting the choice of mathematics-related majors.

To sum up, studies reviewed in this section offer an important message for teaching and learning: self-efficacy instruments can be used as classroom tools by teachers to identify students whose self-efficacy expectations to accomplish a given set of mathematics tasks on a test far exceed or fall short of their achievement. The following subsection offers a chapter summary of studies reviewed.

2.6 Chapter Summary

This literature review is summarised in four subsections as follows. Subsection 2.6.1 deals with students' achievement at the Year 10 level in Australia, together with global comparisons, issues of diversity, such as ability grouping of students for learning mathematics in mixed-ability classrooms, and gender differences in mathematics achievement in Australia and India.

Subsection 2.6.2 summarises studies about students' prior knowledge in learning mathematics as a construct influencing achievement, as well as its operationalisation for assessing learning.

Subsection 2.6.3 summarises a review of studies that provide a theoretical framework for mathematics classroom learning environment, as well as those studies that developed and validated different instruments for measuring classroom learning environment and their association with students' achievement in Australia and India.

Finally, Subsection 2.6.4 summarises studies reviewed about i) social cognitive theory including self-efficacy as a mediator and predictor of achievement and ii) self-efficacy instruments used as classroom tool by the calibration of self-efficacy judgements into efficacy expectancies for measuring students' self-efficacy beliefs.

2.6.1 Students' Mathematics Achievement and Issues of Ability Grouping, and Gender Differences in Australia and India

Many studies (Niemann et al. 2017; OECD, 2016a, b) that analysed PISA results over a period observed a gradual decline in the mean scores of Australian students from an estimated average of 524 points in 2003 to 494 by 2015, which is much below the leading nation, Singapore, with 564 points (Gurria, 2016). Stacey et al. (2015) attributed the decline to the inability of mathematics curricula and classroom practices to ignite the interest of students to learn mathematics.

For India, which was not included in PISA, a National Achievement Survey (NAS) for Class X carried out for the NCERT in India by Sreekanth et al. (2015) in 2014-2015 reported that, in the State of Andhra Pradesh and in the Delhi region where my study was conducted, students' mathematics achievement was low to average. However, the all-India Class X examination results, released by the Central Board of Secondary Education (CBSE), showed that a relatively high percentage of students obtained at least 40% for mathematics, but no details were available about achievement scores by ability groups.

Concerning two major issues of diversity in education (namely, ability grouping by streaming or tracking practices and gender differences in mathematics

achievement), many researchers lamented that there are significant diversity barriers to effective teaching or learning (Askew, 2015; Forgasz, 2010, 2015; Sullivan, 2015). That is, negative effects of streaming and tracking might lead to emphasising certain skills to the detriment of other aspects of mathematics, such as communication, meaning and relevance (Sullivan, 2015). Thus, for meeting the goals of Melbourne Declaration by all Australian governments (ACARA, 2009, 2015), mathematics is best learned and taught in mixed-ability class rooms and when schools adopt classroom practices that offer equally-challenging opportunities to all ability groups without necessarily grouping students by streaming or tracking. As Askew recommends, barriers to effective teaching can be grappled with if schools manage the impact of diversity by offering appropriate individualised learning experiences to students in mixed-ability classrooms. This approach also is consistent with the recommendations that emerged from international research on ability grouping (Hattie, 2009; Law, 2014; NCERT, 2012; Oakes, 1992, 2005; Steenbergen, Makel & Olszewski, 2016).

The second area of research was gender differences in mathematics achievement in Australia (Forgasz, 2010, 2015, Forgasz & Leder, 2015), the United States (Spencer, Steele, & Quinn, 1999) and India (Sreekanth et al., 2012). Hill et al. (2010, p. 38) concluded that "negative stereotypes about girls' and women's abilities in mathematics and science persist despite their considerable gains in participation and performance in these areas during the last few decades".

On the other hand, Else-Quest, Hyde and Linn (2010) meta-analysed the 2003 PISA results for Australia and indicated negligible gender difference in mathematics achievement at the Year-10 level, with a small effect size of 0.06 SD.

2.6.2 Prior knowledge, Its Operationalisation and Students' Achievement

A review of past studies about learning theories supports prior knowledge as a key construct and predictor of students' achievement (Ausubel, 1968, 2000; Dochy, 1992; Schneider, Korkel, & Weinert, 1989; Tobias, 1994). In his assimilation theory, Ausubel (1968) argued that meaningful learning occurs when learners connect new information to the existing knowledge structure, which is characterised by learners' prior knowledge at its core, meaningful material, and learners' intent and ability to use

their prior knowledge. Schneider, Korkel and Weinert (1989) showed that prior knowledge compensated for low aptitude, while high aptitude could not compensate for low prior knowledge. Because individual differences in knowledge are a primary determinant of achievement, students who lack relevant prior knowledge face difficulties in learning new information (Ausubel, 2000; Dochy, 1992; Tobias, 1994).

For operationalising prior knowledge, many cognitivists recognised the notion of conditional knowledge and two types of declarative knowledge (Alexander, Schallert, & Hare, 1991; Paris, Lipson, & Wixson, 1983). Following this, Hailikari, Nevgi and Lindblom-Ylanne (2007) provided a framework for operationalising it in classroom assessment with two types of declarative knowledge (Knowledge of facts and Knowledge of meanings and concepts) and two types of procedural knowledge (Integration of knowledge and Application of knowledge) that describes the nature of content in mathematics tasks.

By assessing prior knowledge according to this framework, numerous researchers investigated plausible relationships between students' prior knowledge and their learning outcomes and acknowledged the significant and positive influence of prior knowledge on learning (Bloom, 1976; Dochy, 1994; Weinert & Helmke, 1998).

Using structural equation modelling, Hailikari, Nevgi and Komulainen (2007, p. 59) reported that "domain-specific prior knowledge was the strongest predictor of student achievement over and above other variables included in the model". The authors concluded that, together with previous study success, it explained 55% of the variance in mathematics achievement.

Also, when the role of prior knowledge in organising students by ability groups for instructional practices was investigated, Hallinan & Sorenson (1987) noted that "the higher a student's standardised test score in mathematics, the higher the likelihood that the student will be assigned to the high-ability group". It is therefore recognised that grouping of students based on their prior knowledge is useful for mathematics instruction within a mixed-ability class.

2.6.3 Theoretical Framework for Mathematics Classroom Learning Environment and Instruments Used for Its Measurement

This review encompassed the seminal work of Vygotsky (1962, 1978) as synthesised by Moll et al. (1980), in which ZPD explains how students learn mathematics at the earlier stages, and how scaffolding aids in knowledge construction that occurs subsequently (Wood, Bruner & Ross, 1976). Interrelated with scaffolding is constructivist theory, which suggests that individual learners construct mental models to understand the world around them. The construction of new knowledge is developed from a state of disequilibrium, when the learner is confronted with an object which could not be assimilated into prior knowledge and thus reconstruction occurred (Harlow, Cummings & Aberasturi, 2006). von Glasersfeld (2000, p. 4) noted that the epistemological belief of a radical constructivist is "one that holds that knowledge is under all circumstances constructed by individual thinkers as an adaptation to their subjective experience"

Brooks (1990, p.69) elaborated that "…the intellectual opportunities the teacher offers the students are carefully constructed invitations that maximize the *possibility* that new conceptual learning will occur". That is, the constructivist teacher can present new information in the context of a problem that introduces disequilibrium, but that student finds interesting and can solve with some effort and support. This process is experienced by students across all grade levels and ability groups.

Another view about mathematics learning is the social perspective, in which mathematics learning is developed in stages by teachers or significant 'others' including peers at school not only by way of intriguing questioning, but also facilitating peer interaction. In his philosophy of social constructivism, Ernest (1991, 1998) argued that mathematics is a language that is socially constructed by learners. Also, from a socio-cultural perspective, learners construct knowledge when teachers and students act as both explainers and listeners, sharing their responsibility for thinking and learning by social interactions in the classroom (Wood & Turner-Vorbeck, 2001)

Regarding the development of instruments, past research reviews recognised the constructivist's approach for assessing students' perceptions of classroom learning environment in different countries (Aldridge, Fraser, Taylor, & Chen, 2000; Fraser, 2002, 2012, 2014; Taylor, Fraser, & Fisher, 1997) and provided compelling evidence about the validity of the WIHIC questionnaire which was extensively adopted in different curriculum domains in different countries for science learning. Aldridge et al.'s (2012) study incorporated these dimensions along with others in developing the COLES that has 11 8-item dimensions, including Personal Relevance, Shared Control, Uncertainty, which also match Six Principles of Learning and Teaching that are followed in Victorian schools. Also, the Differentiation dimension of the classroom learning environment (Fraser, 1990) can be viewed as indirectly connected to the construct of prior knowledge as embedded in the COLES.

Aldridge et al. (2012) validated the COLES with large sample of 2043 Years 11-12 students from 147 classes in nine schools in Western Australia, covering different subject domains. These researchers also examined associations between classroom learning environment dimensions and learning outcomes including students' attitudes and achievement.

For Indian classrooms, Koul (2003, 2009) and Gupta and Fisher (2012) provided some evidence about science learning environments by administering the QTI, the WIHIC and an attitude scale to 1,021 students from 32 science classes of Class IX (Year 9) and Class X (Year 10) in seven different co-educational private schools in the Indian state of Jammu & Kashmir. However, the review indicates limited evidence of COLES having been applied to mathematics and Indian classroom learning environments.

2.6.4 Self-Efficacy as a Mediator and Predictor of Achievement, and Its Calibration

Bandura's (1977) social cognitive theory when applied to classroom context implies that students' achievement can be modelled by reciprocal interactions between students' thoughts and beliefs and classroom environmental dimensions on the one

hand, and students' attention, effort in learning (see Figure 2.1) on the other. Bandura (1986a, p. 25) emphasised that, "what people think, believe, and feel affects how they behave". Thus, in educational settings, self-efficacy is not an isolated construct but rather an integral component of social-cognitive theory because students' social environment can influence both attitudes and behaviours, as evident in classrooms (Schunk & Pajares, 2004; Urdan & Schoenfelder, 2006)

When various researchers investigated the role of self-efficacy beliefs in students' mathematics classroom learning and achievement (Collins, 1982, as cited by Schwarzer, 2014), they used self-efficacy beliefs as a causal factor in human functioning. Perceived self-efficacy exerted a substantial independent effect on performance. Positive attitudes towards mathematics were better predicted by perceived self-efficacy than by actual ability.

Dorman and Adams (2004) reported satisfactory validity for all scales of the WIHIC and the CLES, as well as positive and statistically-significant correlations between academic self-efficacy and each of the ten classrooms learning environment dimensions, although the multiple regression of academic self-efficacy on all ten dimensions indicated a R^2 value of 19.4% only. The self-efficacy effects were statistically nonsignificant for Student Cohesiveness, Teacher Support, Personal Relevance and Shared Control.

In a California-based study which estimated two-level structural models, Fast, Lewis and Bryant (2010) assessed the effect of the perceived classroom environment on mathematics self-efficacy, as well as the effect of mathematics self-efficacy on standardised mathematics test performance, among upper-elementary school students. Students who perceived their classroom environments as more caring, challenging and mastery-oriented had significantly higher levels of mathematics self-efficacy, and higher levels of mathematics self-efficacy positively predicted mathematics performance.

Hailikari, Nevgi and Komulainen (2007, p. 59) concluded that academic self-beliefs strongly correlate with previous study success, with a strong direct influence on prior knowledge test performance, but that students' self-beliefs affected

achievement only indirectly via prior knowledge. The results imply that both prior knowledge and self-beliefs should be considered when examining instructional support issues because they can provide valuable insights about future performance.

But, the above studies did not consider self-efficacy at a task-specific level. Bandura (1980) and Pajares and Miller (1995) made a strong case that self-efficacy should be investigated for individual tasks rather than by global indicators such as confidence in learning (Fennema & Sherman, 1976) because of its behavioural and situational specificity.

My review identified researchers who designed instruments to measure self-efficacy at different levels (viz., mathematics test in the whole domain, or for tasks within the subject domain) and for different purposes (Dorman & Adams, 2004; Fennema & Sherman, 1976, 1977; Hackett, 1985; Midgley & Urdan, 1995; Roeser et al., 1996).

Some other researchers investigated the mediational role of self-efficacy beliefs in quantifying the influence of student-related cognitive factors, including prior knowledge on achievement (Fast et al., 2010; Hackett & Betz, 1989; Randhawa et al., 1993; Siegel et al., 1985) and calibrated self-efficacy judgements into students' self-efficacy expectancies (expected scores) which were then compared with students' achievement scores as awarded by teachers for identifying underachievers, or students who displayed lower efficacy-expectancies than what they achieved (Hackett & Betz, 1989). Bandura (1977, 2006) emphasised that, given incentives to perform, efficacy-expectancies are a major determinant of people's choice of activities, how much effort they expend, and how long they will sustain effort in dealing with stressful situations.

This review, thus, suggests that self-efficacy instruments can be used as classroom tools to identify underachievers for remedial counselling to improve learning outcomes. The next chapter deals with research methods, instruments developed and data- analysis procedures for answering the research questions of my study.

Chapter 3

RESEARCH METHODS

3.1 Introduction

This chapter describes in detail my research methods, instruments, sample and data analyses in eight sections. The research questions and objectives outlined in Subsections 1.5.1 and 1.5.2 guided the selection of the research paradigms of this study, as clarified in Section 3.2.

An overview of research methods is presented in Section 3.3, the instruments developed are considered in Section 3.4 and the sample selection, school background characteristics and demographics are described in Section 3.5.

Section 3.6 offers qualitative analysis of teachers' feedback about how they used their students' prior knowledge in classroom teaching. Also, Section 3.6 elucidates data-analysis procedures for estimating effect sizes for gender differences in prior knowledge, self-efficacy beliefs and achievement, as well as for correlations between i) students' prior knowledge and achievement, ii) classroom learning environment dimensions and achievement and iii) self-efficacy beliefs and achievement, and calibration of students' responses of self-efficacy judgements into expectancies. Later, Section 3.6 details the validation procedures for measuring the construct of classroom learning environment involving a construct validity framework to establish content, face, convergent, discriminant, concurrent and predictive validities, as well as procedures of exploratory and confirmatory factor analyses for extracting the factor structure of the MCOLES.

Section 3.7 presents procedures for i) the Two-level SEM by a variance component analysis of students' achievement and self-efficacy beliefs ii) estimating the mediation model to obtain direct and mediation effects of prior knowledge and classroom learning environment on achievement; iii) the SEM estimation; and iv)

identifying the best SEM variant using evaluation criteria including fit indices. Section 3.8 offers a chapter summary.

3.2 Research Paradigms

A research paradigm is a set of common beliefs and agreements shared between scientists about how problems should be understood and addressed (Kuhn, 1962). According to Guba (1990), research paradigms can be characterised by:

Ontology – What is reality?

Epistemology – How do you know something?

Methodology – How do you go about finding it out?

Methods -- What research tools and techniques used to find out?

As explained below, my research involved a mixed-methods approach. Among others, there are two paradigms relevant for this study, named positivists and constructivist, as described below.

Positivists believe that there is a single reality, which can be measured and known, and therefore they are more likely to use quantitative methods to measure this reality. Constructivists believe that, because there is no single reality or truth, multiple realities need to be interpreted and, therefore, they are more likely to use qualitative methods to interpret them. The interrelationships between other tenets of research paradigms adopted are shown in Table 3.1 along with a description of relevant methods used for addressing each research objective.

The first research question was addressed qualitatively by involving students and teachers in a classroom and seeking feedback from teachers on the use of students' prior knowledge in teaching. This objective was examined with interpretivism or constructivism as a research paradigm involving a qualitative method. Research questions 2 to 5 were addressed using a positivist paradigm in which ontologically a

Table 3.1: Research Paradigms of the Proposed Study using A Mixed-Methods Approach

Research Questions (See Subsection 1.5.1)	Paradigm	Ontology (the reality?)	Epistemology (How I can know the reality)	Theoretical perspective (Approach adopted to know something)	Methodology (How to go about finding out?)	Method (Techniques and tools employed to find out?)
1	Interpretive (Constructivist)	There is no single reality or truth (more realist).	Social constructivism, in which reality needs to be interpreted from the underlying meaning of classroom events and activities of students and teachers.	Interpretivism	Qualitative research	Qualitative: Teachers' reflective Journal (TRJ) notes based on classroom experiences and feedback.
2 - 4	Positivism	There is a single reality or truth.	Reality can be measured and, hence, the focus is on reliable and valid tools to obtain that.	Positivism	Survey research	Quantitative: Sampling, statistical analysis, correlations and factor analyses.
5	Positivism	There is a single reality or truth	Reality can be measured and, hence, the focus is on reliable and valid tools to obtain that	Positivism	Survey research	Quantitative: Two-level structural equation modelling, mediation analysis using hierarchical survey data.

Note: Adapted from Crotty (1998) and Patel (2015).

reality exists, and it is possible to use quantitative techniques and tools in the investigation of truth.

Following the positivist and constructivist paradigms, I employed a mixed-methods approach by combining a qualitative method based on a text analysis of notes and feedback from teachers (Fraser & Tobin, 1991; Merriam, 2009) with a quantitative analysis of data collected by administering surveys to a sample of students and teachers (Creswell & Clark, 2007; Fraenkel, Wallen & Huyan, 2012; Jaeger, 1997

3.3 Overview of Research Methods and Models

This section elucidates the methods adopted as well as research models posited. First, Subsection 3.3.1 provides an overview of different research methods adopted for examining research questions 1 to 4. To address research objective 5(a), Subsection 3.3.2 provides a brief description of the mediation model used for estimating direct and indirect effects of students' prior knowledge and classroom learning environment on achievement, with self-efficacy as mediator. To accomplish research objective 5(b), Subsection 3.3.3 describes the structural equation model (SEM) used for examining the joint influence of students' prior knowledge and classroom learning environment on self-efficacy and achievement.

3.3.1 Research Methods

The first research question involved conducting a qualitative assessment of how teachers used their students' prior knowledge in classroom teaching. To seek teachers' feedback on their classroom experiences, I designed a template of a teachers' reflective journal (TRJ) as explained in detail in Subsection 3.4.2 and qualitatively analysed the TRJ notes collected from all teachers in the sample.

To examine research questions 2 to 4, I proposed a research model as presented schematically in Figure 1.1 of Chapter 1, which has the criterion variable of students' achievement (ACH), the mediator variable of students' self-efficacy beliefs in

achieving (SEI) and the two explanatory variables of students' prior knowledge (PK) and classroom learning environment (CLE).

For examining research objective 2(a), the variables PK and ACH were measured by students' scores on a pretest of prior knowledge and a posttest on achievement, respectively. These tests were specially designed according to the knowledge framework (see Appendix 1) recommended by Hailikari, Nevgi and Lindblom-Ylänne (2007) after consulting teachers on a topic of measurement selected from the Year 10 mathematics curriculum. The data on variables PK and ACH were collected from teachers who administered the pretest and the posttest to their students. Prescores and postscores were analysed by ability groups for boys and girls separately to test gender differences in students' prior knowledge and achievement.

For examining research objectives 2(b) and 2(c), I conducted an analysis of correlations between PK and i) ACH, ii) teaching experience and iii) class size. Improvement (or decline) in students' learning was analysed by examining the difference between students' prior knowledge (prescores) and achievement (postscores) for each ability group.

For measuring the construct of classroom learning environment (CLE) (objectives 3(a) and 3(b) in Subsection 1.5.2), the new Mathematics-related Constructivist Oriented Classroom Learning Survey (MCOLES) was developed as explained in Subsection 3.4.3 and, after teachers administered it to their classes, the MCOLES was validated using students' responses. The factor score coefficients and factor values of the MCOLES were derived by the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) using M*plus* software (Version 7.4). The factor values, thus extracted, were used as a measure of CLE in further analysis of correlations and estimation of the mediation model and the SEM.

For examining the predictive ability of the MCOLES (research objective 3(c) in Subsection 1.5.2), a correlation analysis was conducted between the pairs of variables: (CLE, ACH) and (PK. ACH). The relative importance of different dimensions of classroom learning environment was identified by an analysis of

correlations between students' achievement and MCOLES factors to inform teaching practice for improving students' achievement in the future.

For research objective 3(d), the effect sizes for gender differences in correlations between classroom learning environment and achievement were derived by using the formula recommended by Cohen (1988) and tested for significance by using the Fisher's *r*-to-*z* transformation (Lowry, 2017).

To address research objective 4(a) involving designing a new Mathematics Self-Efficacy Scale (MSES), first, the resource material of mathematics tasks of different types was developed for the same topic of measurement for which students' prior knowledge was assessed. The content validity of the resource material was established by using the criteria of *equity* and *fairness* for the assessment of learning. The face validity of mathematics tasks was justified by feedback from teachers and students (Munby, 1997; Trochim & Donnelly, 2006).

For examining research objective 4(b), students were requested to self-report their self-efficacy judgements about how capable they were to achieve the tasks given in the resource material successfully. The response format included choices on a scale of 1 to 5, where 1 indicates 0% of expected score (i.e., the given task is very difficult to solve) and 5 represents 100% expected score (i.e. the student is very sure about solving it successfully).

Students' self-efficacy judgements were calibrated by converting their self-reported MSES responses into equivalent expected scores, as a measure of their efficacy-expectancies, and expected scores were used to represent SEI (Bandura, 2006). After the MSES was administered, students also provided their solutions to these tasks on the posttest under strict test conditions for their teachers to evaluate. A comparison between students' expected scores and postscores identified underachievers as well as those with lower efficacy-expectancies; this information was used to recommend an appropriate teacher intervention program such as remedial counselling for underachievers.

Furthermore, possible causes of underachievement were investigated using the joint frequency distributions of underachievers in each ability group based on

prescores, as well as in each self-efficacy group based on expected scores, with the aim of examining if low prior knowledge was a cause of underachievement.

For examining research objective 4(c), correlations between PK and SEI and between ACH and SEI were obtained using students' prescores, postscores and expected scores and, a correlation analysis was conducted for the overall sample, each school and each class.

To accomplish research objective 4(d), gender differences in efficacy-expectancies and gender differences in correlations between efficacy-expectancies and achievement were examined and tested for significance for each school, using online tools recommended by (Lowry, 2017).

3.3.2 Mediation Model

For addressing research objective 5(a), self-efficacy was employed as a mediator, following MacKinnon et al. (2007), in the mediation model depicted by equations (3.1) and (3.3) below:

$$ACH = a_1 + c\,X + e_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3.1)$$

where c is the direct effect of X on ACH, a_1 is the intercept of ACH, and e_1 is the error term.

$$ACH = a_2 + c'\,X + b\,SEI + e_2 \ldots\ldots\ldots \quad (3.2)$$

$$SEI = a_3 + d\,X + e_3 \qquad \ldots\ldots\ldots \quad (3.3)$$

where a_2 and e_2 are the intercept and error terms of the regression of ACH with SEI as the mediator variable; a_3 and e_3 are those of the regression of SEI; and X is an independent variable such as students' prior knowledge (PK) or classroom learning environment (CLE), whose effects are mediated through SEI on ACH.

An estimate of c' gives the effect of X on ACH adjusted for mediation (MacKinnon, 2007; MacKinnon, Fairchild, & Fritz, 2007), and indirect (or mediated) effects were obtained from the product (b.d) using the estimates of b and d as described in detail in Subsection 3.7.2.

3.3.3 Structural Equation Model (SEM)

To address research objective 5(b), a SEM was hypothesised (see Figure 1.1) based on a theoretical rationale outlined in Subsection 1.4.2, following Bandura (1980) and Urdan and Schoenfelder (2006), who argue that it is important to embrace the social-cognitive view of student motivation and understand that altering controllable factors in the classroom environment could enhance students' motivation towards learning.

Accordingly, it was posited that students' prior knowledge, jointly with different classroom learning environment dimensions (e.g., differentiation) directly influences students' achievement positively, and that these influences are mediated through self-efficacy beliefs.

Thus, the following three hypotheses were proposed for the SEM depicted by equations (3.4) to (3.6);

Students' prior knowledge (PK) jointly with the classroom learning environment (CLE) influences students' achievement (ACH) directly and positively, as shown mathematically by the equation (3.4);

Self-efficacy beliefs in achieving (SEI) is also positively influenced by PK and CLE, as shown by the equation (3.5);

Self-efficacy (SEI) mediates the effects of PK and CLE through its influence on students' achievement (ACH) positively, as shown in the equation (3.6).

$$ACH = f(CLE, PK) \quad \ldots\ldots \quad (3.4)$$

$$SEI = f(CLE, PK) \quad \ldots\ldots \quad (3.5)$$

$$ACH = f(SEI) \quad \ldots\ldots \quad (3.6)$$

Because the data on variables ACH and SEI were hierarchical, the variance components of WITHIN Class and BETWEEN Classes were analysed in Subsection 3.7.1 to justify the choice of modelling of Two-level SEM. The next section describes the instruments used for data collection.

3.4 Instruments

This section describes the development of four instruments used for data collection: i) a pretest of mathematics tasks for assessing students' prior knowledge required for learning a selected topic, which is explained in Subsection 3.4.1; ii) Teachers' reflective Journal for qualitative assessment of how teachers used their students' prior knowledge in classroom teaching in Subsection 3.4.2; iii) the new Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) for seeking students' perceptions of classroom learning environment, in Subsection 3.4.3; and iv) the new Mathematics Self-Efficacy Scale (MSES) to seek students' perceptions of self-efficacy beliefs in successfully achieving tasks assigned in Subsection 3.4.4.

3.4.1 Assessment of Prior Knowledge Required for Learning

For designing the pretest, the researcher first consulted the mathematics curricula prescribed for Years 9-10 in Australia and India, and then the participating teachers, who selected a topic of measurement from the Year 10 mathematics curriculum in Australia, or the topic of *surface areas and volumes of solids* from the Class X Indian curriculum. The mathematics curricula offer equally challenging learning opportunities to students of all levels of ability in Australia (ACARA, 2009, 2012) and India (NCERT, 2015). As Tables 3.2 shows, there is little difference in the curriculum content for the Years 9 and 10, which made it feasible to design a common pretest to assess students' prior knowledge (PK) in terms of their prescores. Thus, the researcher and teachers identified four types of tasks for inclusion in the pretest shown in Appendix 2:

D1: Declarative knowledge of concepts and meanings

D2: Declarative knowledge of facts

P1: Procedural knowledge of integration of concepts, meanings and facts

P2: Procedural knowledge of problem solving applications.

Table 3.2: Comparisons of Mathematics Curriculum between Australia and India for the Competencies of Measurement and Geometry at Years 9 - 10 Levels

Country	Year Level	Content Description
Australia	Year 9	Using units of measurement, calculate the areas of composite shapes (VCMMG312).
		Calculate the surface area and volume of cylinders and solve related problems (VCMMG313).
		Solve problems involving the surface area and volume of right prisms (VCMMG314).
		Using Pythagoras theorem, investigate its application to solving simple problems involving right angled triangles (VCMMG318).
	Year 10	Using units of measurement, solve problems involving surface area and volume for a range of prisms, cylinders and composite solids (VCMMG343).
India	Year 9 (Class IX)	Surface areas of a range of solids viz., two dimensional regular shapes, square, rectangle, circle, trapezium, cylinder, cone etc.
		Volumes of different solids using formulae for prisms, viz., cubes, rectangular prisms. Cones, cylinder, and sphere; Pythagoras theorem and applications.
	Year 10 (Class X)	Surface areas of composite shapes, and volumes of solids involving composite shapes of 3 dimensional objects *viz.*, rocket, tub, a sharpened pencil, frustum etc.

Note: Adapted from http://victoriancurriculum.vcaa.vic.edu.au and http://epathshala.nic.in

For example, to learn how to measure the surface areas and volumes of three-dimensional objects such as cubes, cylinders, spheres and cones, etc. at the Year 10 level, students should have prior knowledge of geometrical shapes and the formulae for areas of two-dimensional objects such as square, rectangle, circle and sector, etc. that are usually taught during the previous year (i.e. Year 9 in Australia or Class IX in India). Thus, the underlying mathematics tasks were designed on the pretest and classified by task-types D1, D2, P1 and P2; example items are shown in Figure 3.1.

3.4.2 Teachers' Reflective Journal (TRJ) on the Use of Students' Prior Knowledge in Classroom Teaching

To accomplish the first research objective, all participating teachers were requested to make use of students' responses to the pretest of prior knowledge required

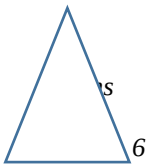
A sample item for D1: Declarative knowledge of concepts and meanings is:
Q. Match the names with the correct geometrical shapes below.

Geometrical shape *Write the correct name from the list given*

(Parallelogram, Rectangle)

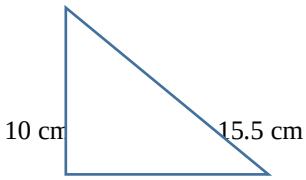
A sample item for D2: Declarative knowledge of facts is*:*

Q. Write an appropriate formula. for finding the area of the following geometrical shape (isosceles triangle) which has two equal sides of length 10 cms each, and a base of 6 cms



A sample item for P1: Procedural knowledge of integration of concepts, meanings and facts is:

Q. Work out the perimeter of the right-angled triangle, given the length of one side is 10 cm and the hypotenuse is 15.5 cms as shown below.



A sample item for P2: Procedural knowledge of problem solving applications is.

Q. Suppose the dimensions of your mathematics text book are given by
Length(L)= 16cm; Width (W) = 12 cms and Height (H) = 8 cms
Work out its total surface area, and volume.

Figure 3.1: Sample Items of Four Task Types D1, D2 , P1 and P2 for Assessing Prior Knowledge Required for Learning a Topic of Measurement at the Year 10 Level

for learning (prescores) to divide their students into three ability groups, and supplement classroom teaching by including suitable tasks in their individual teaching plans for addressing the cognitive needs of ability groups in mixed-ability classes.

They were also requested to complete a template of Teachers' reflective Journal (TRJ) as shown in Appendix 2 by recording their experiences of using students' prior knowledge, and to share their experiences with the researcher at the end of teaching the selected topic. The objective of the TRJ was clearly stated as: "To be able to use students' prior knowledge for teaching a mathematics topic of Year 10: Measurement". TRJ posed four different questions for seeking responses. The first question involved teachers clarifying the learning intentions for their class(es); the second question focused on a brief description of teaching tasks used in the class; and the third question sought teachers' clarification of what divide of questions were raised in the class to motivate participation of students of low-, medium- and high-ability groups. The fourth question in the TRJ related to how teachers facilitated questioning in the class and encouraged students to ask questions whenever students wanted clarifications to help them to solve tasks or understand new ideas, etc. This might inform how new knowledge is possibly constructed by social interactions with others and peers in the classroom and help the researcher to interpret how prior knowledge influences students' understanding of new ideas in mathematics learning across all ability groups. The last item in the TRJ was an open-ended question about other observations that teachers made in improving students' learning.

3.4.3 Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES)

For developing the MCOLES, the literature review of mathematics learning theories and learning environments research in Section 2.4 recognised the importance of the Constructivist Oriented Learning Environment Survey (COLES) which was applied to the science domain (Aldridge et al., 2012; Fraser, 2014), as well as its suitability for mathematics classrooms after making some modifications to items for this study. The COLES is underpinned by constructivist theory (von Glasersfeld, 2000), Ernest's (1991, 1998) philosophy of social constructivism in mathematics learning, and Wood and Turner-Vorbeck's (2001) socio-cultural perspective for analysing the roles that teachers and students play in mathematics classrooms, as they

share their responsibility for thinking and learning by social interactions in the classroom.

Also, embedded in the Six Principles of Learning and Teaching adopted by the Department of Education and Early Childhood Development (DEECD) in the State Government of Victoria (2012) are Moos' (1974) three psychosocial environment dimensions:

- Relationship dimension (which identifies the nature and intensity of personal relationships within the environment and assesses the extent to which people are involved in the environment, support and help each other
- Personal Development dimension (which assesses basic directions along which personal growth and self enhancement tend to occur)
- System Maintenance and Change dimension (which involves the extent to which the environment is orderly, clear in expectations, maintains control and is responsive to change) (Aldridge & Fraser, 2008)

For example, as shown in Table 2.2 of Chapter 2, the principle that 'Learning connects strongly with communities and practice beyond the classroom' can be mapped to Moos' notion of Personal Development, which maps to the Personal Relevance dimension of the COLES.

Furthermore, the Differentiation dimension of the COLES is connected to the Moos' System Maintenance and Change dimension, which recognises the importance of prior knowledge of participants in measuring classroom learning environment.

Extensive literature reviews by Fraser (2012, 2014) provided limited evidence that the COLES and other related instruments have been used successfully in different curriculum domains and countries.

Aldridge et al. (2012) incorporated these dimensions along with others in developing the 88-item COLES, which was validated with large sample of 2043 Years 11-12 students from 147 classes of in nine schools in Western Australia, covering different subject domains in general.

To establish the content validity of the MCOLES, all 88 items in 11 dimensions of COLES were closely examined, and necessary changes were made following theories consistent with the principles of social constructivism for mathematics learning, as well as DEECD's six principles of learning and teaching (Aldridge et al., 2012; State Government of Victoria, 2012).

The MCOLES has seven dimensions. Out of them, Teacher Support, Involvement, Equity and Differentiation dimensions were adapted from the COLES with minor modification to ensure that the item wording was familiar to students. For example, "the teacher is interested in my problems" (COLES) was changed to "my mathematics teacher is interested in my problems" (MCOLES). The remaining three dimensions of the MCOLES were constructed as follows.

Originally, in the COLES, the two separate dimensions of Student Cohesiveness and Personal Relevance were closely interrelated. All 16 items of these two dimensions were divided and rewritten as eight coherent and concise items under the new MCOLES dimension of Student Cohesiveness and Personal Relevance, mainly for brevity and more meaningful organisation of items to reduce student fatigue, as shown in Table 3.3. All of these items map to the Relationship dimension of Moos (1974, 1979).

Similarly, all 16 items of the COLES dimensions of Cooperation and Task Orientation were divided into eight pairs of interrelated items and rewritten for the MCOLES dimension of Task Orientation by Cooperation. All 16 items in the other COLES dimensions of Formative Assessment and Clarity of Assessment Criteria, were combined pairwise and reduced to eight items in the new MCOLES dimension of Clarity of Assessment Criteria & Feedback. Lastly, the COLES dimension of Young Adult Ethos, intended for mature youth in Years 11 and 12, was dropped because of its limited relevance for Year 10 students. In summary, the 56 item-MCOLES required less response time than the 88-item COLES. Appendix 8 provides a copy of all 56 items of the MCOLES.

Table 3.3: Comparison between COLES and MCOLES: Dimensions (Scales) and Sample Items

Dimensions and Items of COLES	Dimensions and Items of MCOLES
Student Cohesiveness	Student Cohesiveness and Personal Relevance
I make friends among students in this class. Members of the class are my friends.	I make friends with many students in my mathematics class and many of them are my friends.
Personal Relevance	
I draw on past experiences to help me in this class. I apply my past experiences to the work in this class.	I draw on past experiences and apply to the work in my mathematics class.
Task Orientation	Task Orientation and Cooperation
Getting a certain amount of work done is important to me. I know how much work I have to do.	I know getting a certain amount of mathematics work done is important and how much mathematics work I have to do.
Cooperation	
I cooperate with other students when doing assignment work. I learn from other students in this class.	I cooperate with other students and learn from them while doing mathematics assignment work in the class.

(contd...)

Table 3.3 **(**contd..2**):** Comparison between COLES and MCOLES: Dimensions (Scales) and Sample Items

Dimensions and Items of COLES	Dimensions and Items of MCOLES
Clarity of Criteria of Assessment	Clarity of Assessment Criteria & Feedback
I am aware of which activities and tasks are used to assess my performance.	For the mathematics assessment, the criteria are clear to me as I know which activities and tasks are used to assess my performance.
The assessment criteria are clear to me.	
Formative Assessment	
I use feedback from assessment tasks to improve learning.	For improving my mathematics learning, I use feedback from my assessment tasks and understand their links with classroom activities.
There is a link between classroom activities and assessment tasks.	

At the second stage of the development of the MCOLES, experienced mathematics teachers were asked to assess the comprehensibility, clarity and accuracy of items for each scale. 11 teachers who participated in the study and reviewed the initial versions of the instruments had more than ten years of teaching experience in secondary mathematics classes, with two of them being expert teachers having more than 30 years of teaching experience. They administered the pilot version of MCOLES to a subsample of 20 low-, 20 medium- and 15 high-ability students, which accounted for over 10% of the main sample. This was followed by personal interviews and feedback from teachers. Some students from the Australian school made useful observations about certain items and even proposed modifications that led to the redrafting of those items, which contributed to a better understanding and improvement of the MCOLES

3.4.4 Mathematics Self-Efficacy Scale (MSES)

For measuring the construct of self-efficacy beliefs in achieving (research objective 4(a) of Subsection 1.5.2), the new Mathematics Self-Efficacy Scale (MSES) was developed with resource material of four different task-types, D1, D2, P1 and P2, designed for the mathematics topic of measurement at the Year 10 level, according to the knowledge framework recommended by Hailikari, Nevgi and Lindblom-Ylänne (2007). Students' prior knowledge was assessed on the same topic of measurement as discussed in Subsection 3.4.1. The test content in the resource material for the MSES was also the posttest material which was trialled initially with Year 10 mathematics teachers. Based on a review, the tasks were classified in an increasing order of difficulty from the students' perspective, which led to creating the two further categories of P3: Procedural knowledge for solving mixed tasks and P4: Procedural knowledge needed for higher-order thinking and solving challenging tasks.

Because the survey was first administered in an Australian school, the posttest tasks were first designed according to the Australian Mathematics Curriculum (ACARA, 2015). Later, some of these tasks were modified to suit the Indian curriculum (NCERT, 2015) delivered by Indian school teachers of Classes 3 to 16.

Both versions of the posttest consisted of 31 tasks which were trialled with all teachers who participated in the study to ensure that teachers delivered the underlying curriculum covering concepts, meanings and application tasks.

Also, Table A9.1 in Appendix 9 shows the classification of all tasks on the posttest and their corresponding weights (scores) assigned for Class 1 of School 1 and Classes 3 to 16 in other schools, separately. When scores are aggregated over all questions of a given task type, as shown in Table A9.2 in Appendix 9, the proportion of all declarative type tasks in the posttest was 25 % for Class 1 and 20 % for other classes. For procedural knowledge. types of P1 – P4, the combined weight for class 1 was 75%, which is slightly less than for other classes (80%). Thus, at the aggregate level, there was a little difference in weights assigned to tasks on the posttest between Class1 and other classes in the sample, which supports the use of students' achievement data based on the pooled sample of six schools in this study.

The MSES, designed according to a guide recommended by Bandura (2006), had a response scale of 1 to 5 and was structured to mirror the tasks set in the resource material, as shown in Appendix 5 for School 1 and in Appendix 7 for other schools. The survey was administered after teachers explained to their respective classes how to give responses as intended and clarified any doubts by means of sample questions as given below. For example, the items at Q. No 1(a) and 1(b) of the resource material (Appendix 4) read as:

Q.No 1 Change the units in the following as indicated:

a) *8 m = --- cm*
b) *6 km = --- m*

Then students were guided to identify the corresponding survey item with the same serial numbers, Q. No 1(a) and 1(b) in the MSES (Appendix 5), and to give a suitable response to the same items. Depending on the student's understanding and self-assessment of ability to answer the given questions, he/she would circle an appropriate response on a scale of 1 to 5. Teachers guided students in how to refer to a specific item serially numbered in the MSES after having read it from the resource material, and in how to give a response to the same item by making a circle around an

appropriate choice provided. For instance, if the student thought that he/she knew the correct answer and was very sure, he/she would choose 5, which corresponds to an expected score of 100% (see Figure 3.2). If unsure, he/she would choose any other relevant choice that reflects his/her self-assessment of ability to do so. Soon after the survey was completed, all students were provided a 15-minute break to relieve survey fatigue. Later, students responded to the posttest under strict test conditions. Their teachers scored and shared the postscores with the researcher.

Responses		Expected Scores
1	If you don't know how to solve that problem	0 %
2	If you're unsure, but can try	25 %
3	If you think 'Maybe'	50 %
4	If you're sure you can solve the problem	75 %
5	If you're very sure that you can do it successfully	100 %

Figure 3.2: Calibration of Self-Efficacy Judgements to Convert Responses into Expected Scores

3.5 Sample Selection, School Background Characteristics and Demographics

This section presents details of sample features in three subsections. Subsection 3.5.1 describes sample selection process, while Subsection 3.5.2 describes important background characteristics of schools in the sample. Subsection 3.5.3 provides details of demographics including the sex ratio of participating students and teachers.

3.5.1 Sample Selection

A rationale for selecting Year 10 students from an Australian school and four Indian schools in this study was discussed in Section 1.2. The sample, chosen carefully in two stages, was purposive to meet the objectives of this study.

At stage 1, letters of invitation were sent to the principals of all government secondary schools covering Year 10 in the Western Region of the State of Victoria and to the principals of selected schools covering Class X in India. All schools whose

teachers accepted and volunteered to take part in my study were included in the sample.

At stage 2, letters of invitation were sent to the parents of Year 10 mathematics students, and their mathematics teachers at participating schools; all students whose parents provided written consent (complying with Curtin university ethical guidelines) were included on a come-first basis until the sample reached a desirable size.

The participants in the present study included 531 students and 11 teachers who taught them mathematics at Year 10 level in Australia (or Class X in India). Because all students who were sent letters of invitation accepted to participate in the study and were present in their classrooms during the surveys, the response rate of students and teachers was 100%.

3.5.2 School Background Characteristics

The sample had six typical co-educational schools with 15 Year 10 mathematics classes. Each school was assigned an ID from 1 to 6. School 1 is a rural school in a Western suburb of Victoria, Australia. Schools 2 to 5 are the government-aided, co-educational schools that cater for the needs of predominantly local community groups spread over different suburbs of the National Capital Region of Delhi, India, and run by a private society since 1948 (see Figure 3.3 for geographic locations in India).

In contrast, School 6 is a boarding school with separate hostel facilities for boys and girls, who also are provided with tutoring support from a homework club and are closely monitored by a dedicated team of resident teachers with a duty of care. This school is globally known for imparting high-quality education and disciplined learning, with a focus on the character development of each student and guided by the principles of Bhagavan *Sri Satya Sai Baba*, whose style of teaching-learning typically resembles the *Guru-Sishya* relationship of the age-old Indian tradition in which the Sishya (student) receives knowledge and blessings from the Guru (teacher).

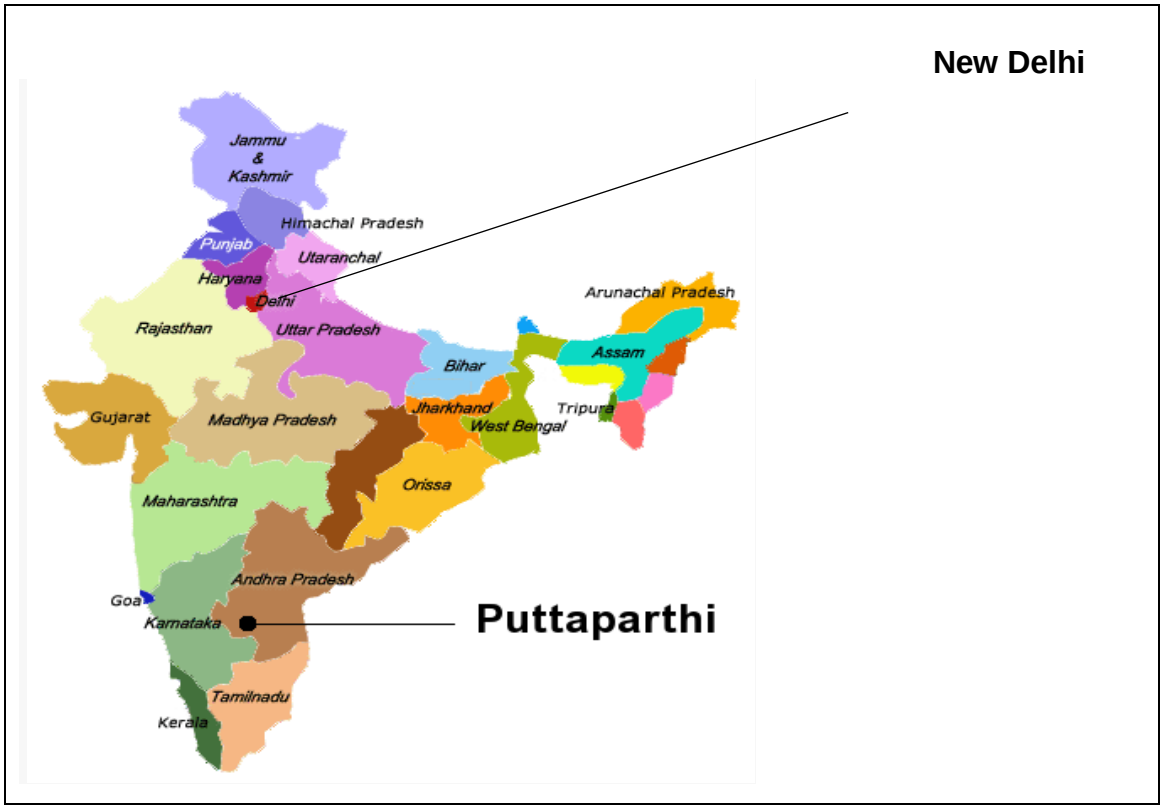


Figure 3.3: Map of India Indicating the Locations of Schools, Puttaparthi and New Delhi

3.5.3 Demographics

For the overall sample of 531 students, the ratio of boys to girls was 1.6:1. The male-female ratio however, varied considerably for different classes. In School 6, Classes 14 and 15 had all boys, while Class 16 had all girls. For other co-educational classes, the sex ratio varied. In Class 5, there were equal number of boys and girls, while Class 12 had more girls than boys.

All mathematics teachers were experienced, with wide-ranging teaching experience from 10 to 32 years for female teachers and from 15 to 40 years for male teachers. Most students (10 out of 15 classes) in the sample were taught by female teachers. The following section describes procedures adopted for data collection.

3.6 Data Collection, Missing Data and Data Analyses

This section provides details of data collection, how missing data were imputed, and data-analysis procedures used to accomplish my research objectives in five subsections. Subsection 3.6.1 provides details of all data collected, how missing data were imputed after coding and data entry, and how accuracy and internal consistency were checked, and the resultant sample size available for analyses.

Subsection 3.6.2 describes the method used for qualitative investigation of how teachers utilised data on their students' prior knowledge in classroom teaching, which involves procedures of analysis of teachers' feedback based on notes from teachers' reflective journal (TRJ) (research objective 1).

Subsection 3.6.3 explains data-analysis procedures adopted for measuring students' prior knowledge by each ability group and by gender, and for testing gender differences statistically (research objectives 2(a) and 2(b)).

Subsection 3.6.4 provides procedures adopted for validating the MCOLES and its associations with students' prior knowledge and achievement to address research objectives 3(a) to 3(d).

Subsection 3.6.5 explains the calibration of self-efficacy judgements into efficacy-expectancies (research objective 4(b)) for measuring students' self-efficacy beliefs (SEI), investigating associations of SEI with PK, CLE and ACH variables (research objective 4(c)), and exploring gender differences in efficacy-expectancies (SEI) (research objectives 4(d)).

3.6.1 Data Collection, Missing Data and Sample Size

Collection of data on prior knowledge involved a pretest of tasks (see Subsection 3.4.1) for the topic of measurement related to the Year 9 mathematics curriculum. Teachers who administered this test scored it and shared the prescores with the researcher.

For the qualitative part of this study, teachers were requested to record their experiences of using mathematics tasks in classroom teaching in the specially-designed TRJ explained in Subsection 3.4.2, mainly to interpret how they used tasks to cater for the individual needs of students who were divided into ability groups based on their prior knowledge.

To assess the construct of classroom learning environment (CLE), students' responses to various items of the MCOLES were measured on a frequency scale from 1 to 5 (1 for Almost Never to 5 for Almost Always). The MCOLES was administered in all classes after teachers taught a chosen topic, so that students could reflect on their perceptions of what happened in their classrooms when responding to the MCOLES.

Similarly, to assess self-efficacy beliefs in achieving (SEI), a set of mathematics tasks were selected on the chosen topic of measurement in the resource material (posttest). At the end of teaching this topic, in a preamble to the survey, teachers informed students about the assessment criteria for the posttest on this topic, and about the MSES, as explained in Subsection 3.4.4. At this stage, students were not to answer the test questions but only indicate their ability to answer them successfully. Later students responded to the posttest under strict test conditions, teachers scored it and shared postscores with the researcher, who used postscores as a measure of students' achievement (ACH).

This study also calibrated students' self-efficacy judgements, as suggested by Bandura (1980), by converting students' MSES responses into expected scores which reflect their self-efficacy-expectancies. Expected scores on the posttest were used as a measure of students' self-efficacy beliefs in achieving given tasks (SEI).

Before commencing data analysis, data clean-up was undertaken as recommended by Alreck and Settle (1995). This process involved checking each student record for data-entry errors and/or missing or implausible data.

A detailed data review by another trustworthy person identified 65 incomplete records for the MCOLES and 85 incomplete records for the MSES. To impute the missing responses with randomly-generated data, I adopted the statistical technique of Missing Completely At Random (MCAR) (Little & Rubin, 2002; Muthen & Muthen,

2008) by using the M*plus* software (Version 7.4). According to Enders (2010, p. 7): "The formal definition of the MCAR requires that the probability of missing data on a variable Y is unrelated to other measured variables and is unrelated to the values of Y itself". Use of imputed data generated by the MCAR technique implies that the observed data points are a simple random sample of the scores available if the data had been complete. The data analysis finally utilised valid responses to the MCOLES from 511 students including 198 girls and 313 boys and, to the MSES, from 491 students including 184 girls and 307boys.

After imputing data, it was found that some students complied with the MCOLES and / or the MSES but did not take part in the pretest and / or posttests and, hence, were dropped from the analysis.

When only those students who completed all requirements for the MCOLES and the MSES, as well as the pretest and the posttest, were considered, the sample size was reduced to only 464 students (179 girls and 285 boys) who were available for model estimation.

3.6.2 Analysis of Teachers' Feedback and TRJ Notes

To address the first research objective of how teachers used students' prior knowledge in classroom teaching, teachers selected the topic of measurement from the Year 10 mathematics curriculum, designed and conducted a pretest of prior knowledge required for learning this topic, and used prescores to divide students into three ability groups for teaching the selected topic.

Following Sullivan's (2015) study, teachers were requested to design suitably challenging tasks for use in the classroom teaching to students of all levels of ability. Then, they were requested to address four questions raised in TRJ and record their classroom experiences by completing a template (see Appendix 3). To address the first research objective, TRJ notes reflecting teachers' answers to these four questions were proposed to be tabulated and reported for convenience of analysis, as shown in Table 4.2, for analysis using simple visual inspection and interpretation.

3.6.3 Measurement of Students' Prior Knowledge and Achievement by Ability Groups and Gender

Addressing research objectives 2(a), 2(b) and 2(c) of Subsection 1.5.2 required data on students' prior knowledge (PK) and achievement (ACH), which involved students' scores on the pretest and posttest on the topic of measurement, as well as data on classroom variables, teaching experience and class size.

To test the statistical significance of gender differences in prior knowledge and achievement, first, I worked out the respective means and standard deviations by ability groups and gender by using the SPSS (Version 24.0), and then used the hypothesis testing procedures (Snedecor & Cochran, 1989) given below.

The null hypothesis H_0 states that there is no difference between the two population means, and the alternative hypothesis H_a states that there is:

H_0: $m_1 = m_2$, H_a: $m_1 \neq m_2$

The test statistic is given by:

$t = (m_1 - m_2) / \text{Std.Error}$

where Std.Error $= \sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]}$, given that s_1 and s_2 are the standard deviations, and n_1 and n_2 are the sample sizes of boys and girls, respectively. The critical values from the Student's *t*-distribution were compared with their respective computed values of the *t*-statistics at the 5% (or 10%) level at the appropriate degrees of freedom (df), under the assumption that populations were normal. For unequal variances, df is given by

$$(s_1^2/n_1 + s_2^2/n_2)^2 / \{[(s_1^2 / n_1)^2 / (n_1 - 1)] + [(s_2^2 / n_2)^2 / (n_2 - 1)]\}.$$

Furthermore, the magnitudes of gender differences were obtained using the effect size (*d*), computed according to Cohen (1988) using the formula,

$$d = (m_1 - m_2) / \sigma_{pooled},$$

where $\sigma_{pooled} = \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$ after replacing s_1 by σ_1 and s_2 by σ_2.

d can also be derived from the *t*-value of the differences between the two groups (Rosenthal & Rosnow, 1991).

3.6.4 Validation of MCOLES and Its Associations with Achievement

For accomplishing research objective 3(a), the MCOLES was developed and validated following Trochim and Donnelly's (2006) framework, which also has been widely used by others (e.g. Velayutham, 2012). It suggests that a construct must fulfil both *translation* and *criterion-related validity* requirements. Translation validity includes i) content validity in that the items underlying the construct must be theoretically sound and ii) face validity in that items are clearly interpretable by students and teachers as intended by theory. Criterion-related validity includes i) convergent validity in that items within a scale are highly correlated with each other and ii) discriminant validity in that items from different scales are not highly correlated. Furthermore, the construct in a questionnaire must have concurrent validity in that it can distinguish between groups of participants as anticipated on theoretical grounds. Finally, a construct must also have predictive validity in that it must be able to predict what it should, theoretically.

Translation validity was investigated based on sound theories and feedback from a pilot survey involving students and teachers. Convergent validity was established using Cronbach's alpha coefficient to check scale consistency. A cut-off value for *r* of 0.4 was used to retain an item within a scale. Thus, items were excluded if their inter-item correlation was less than 0.4 with others within a scale. Also, items that were negatively correlated with others were regrouped with other suitable items to form a subgroup if they satisfied the cut-off value.

Discriminant validity was investigated by bivariate and partial correlation analyses of item means, while concurrent validity was examined by one-way ANOVA (Tabachnick & Fidell, 2013).

Factors were extracted by exploratory factor analysis (EFA) using the SPSS software (research objective 3(b)). For the determination of the number of factors, factor analyses of the MCOLES were conducted involving:

i) principal axis factoring with direct oblimin or the varimax rotation for the factorability of data (Hubbard & Allen, 1987; Schmitt, 2011)
ii) scree plots of eigenvalues
iii) eigenvalue comparison by the Monte Carlo Parallel Analysis (Watkins, 2000).

Removal of items that were weakly intercorrelated ($r < 0.4$) in each scale and regrouping of items resulted in rearrangement of dimensions of the MCOLES, as identified in the EFA. Then, confirmatory factor analysis (CFA) was applied to all items of dimensions (scales) jointly, by applying the M*plus* software, and maximum likelihood estimates of factor loadings, factor score coefficients and factor values were obtained from the CFA of MCOLES responses.

The latent structure of the construct of classroom learning environment (CLE), characterised by first-order factors, was further examined for any possible intercorrelations among these fist-order factors both on theoretical and empirical grounds in search of an acceptable parsimonious solution. If the first-order factors are not theoretically related, and if correlations between them are not statistically significant, there would be no justification in pursuing higher-order factor analysis whereas, if there exists a high correlation amongst them, then the second-order CFA could be applied, as it should be on theoretical basis (Brown, 2014).

For checking the predictive validity of the MCOLES (research objective 3(c) of Subsection 1.5.2), correlations were calculated between i) CLE and ACH, and ii) CLE and PK, using the standard formula of Pearson correlation coefficient (Kendall et al., 1973) and the SPSS software program.

To test the significance of correlation coefficients *(ρ)*, the null-hypothesis H_0 and the alternate hypothesis by H_A are given by:

$H_0: \rho = 0$

$H_A: \rho \neq 0$

To test this hypothesis, a *t*-value was computed for each estimated correlation coefficient (*r*) using the formula:

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

where *n* is the number of pairs of variables X and Y, with X being students' prior knowledge and Y being their achievement. To test if a sample correlation is significantly different from zero, the respective *t*-value is compared with the critical value from the Student's *t*-distribution at the 5% (or 10%) level at the degrees of freedom (*n* -2).

Furthermore, partial correlations were also investigated using the SPSS because bivariate correlations between X and Y could be biased upwards, if there is a co-covariable C which is also correlated with Y. In this case, the variable C confounds the influence of X on Y (Fisher, 1924; Kendall et al., 1973). In the present context, X=PK. Y=ACH and C=CLE. Thus, the bivariate correlation between PK and ACH could be biased upwards because classroom learning environment (CLE) is also correlated with achievement (ACH).

For examining research objective 3(d) of Subsection 1.5.2, gender differences in correlation (r_b for boys and r_g for girls) were derived. To test the statistical significance of gender differences in correlations (r_b – r_g), the Fisher's *r*-to-*z* transformation was applied. First *z*-score values of correlation coefficients and their difference (r_b – r_g) were derived and, later, *p*-values for hypothesis testing were examined using online tools (Lowry, 2017).

3.6.5 Calibration of Self-Efficacy Judgements into Efficacy-Expectancies and Their Associations with Achievement

As explained in Subsection 3.4.4, the new MSES was developed for assessing students' perceptions of self-efficacy beliefs (research objective 4(a)). Subsection 3.6.5 explains data-analysis procedures adopted for calibration of students' self-efficacy judgements by converting MSES responses into expected scores (research objective 4(b)), as well as data procedures for identifying underachievers (research objective 4(c)) and gender differences in self-efficacy expectancies (research objective 4(d) of Subsection 1.5.2)

First, to meet research objective 4(b), deviations between students' expected scores and achievement (postscores) were calculated as a percentage of postscores, after allowing for an error of ± 5% to consider the limitation of continuity on the MSES response scale, and then used to divide students into the following three categories, using the EXCEL spreadsheet:

- Underachievers (or students with higher efficacy-expectancy) whose expected scores were more than the postscores, with the deviation being positive and more than 5%.
- Students whose expected scores matched their postscores, with the deviation being in the range of ± 5%
- Students with lower efficacy-expectancies whose expected scores were less than postscores, with the deviation being negative, but more than 5% in magnitude.

To identify possible causes of underachievement, underachievers were further divide into three ability groups according to their prior knowledge or prescores:

0 - 40% (Low-ability)

40 - 60% (Medium-ability)

60 - 100% (High-ability)

The joint frequency distribution of efficacy-expectancies and achievement was analysed to investigate any pattern of relationship between prior knowledge and achievement.

Similarly, underachievers were also segregated into three self-efficacy groups based on their expected scores, and the corresponding joint frequency distribution with achievement was obtained to analyse the pattern of underachievers with high-self-efficacy for boys and girls, separately. Students who underachieved because of low prior knowledge or high self-efficacy were identified for an appropriate remedial teacher intervention program.

For accomplishing research objectives 4(c) and 4(d) of Subsection 1.5.2, correlations between efficacy-expectancies (expected scores) and achievement (postscores) were analysed and effect sizes (*d*) were calculated for efficacy-expectancies following Cohen (1988). Gender differences in correlations were also examined using the Fisher's *r*-to-*z* transformation for each school and the overall sample by using similar data-analysis procedures mentioned in Subsections 3.6.3.

The next section addresses data-analysis procedures adopted for meeting research objectives 5(a) and 5(d) of Subsection 1.5.2.

3.7 Model Estimation and Evaluation of SEM Variants

This section addresses research objectives 5(a) and 5(b) of Subsection 1.5.2 in four subsections by providing procedures and methods that involved:

i) analysis of variance components of hierarchical data on the dependent variables, ACH and SEI, for WITHIN Class and BETWEEN Classes, which is explained in Subsections 3.7.1
ii) a mediation model as proposed in Subsection 3.3.2 to estimate the direct and mediated effects of explanatory variables, PK and CLE, on the criterion variable, ACH, using the mediator variable, SEI (research objective 5(a), which is described in Subsection 3.7.2
iii) the Structural Equation Model (SEM) as proposed in Subsection 3.3.3 for estimating the joint influence of explanatory variables, PK and CLE, on ACH in Subsection 3.7.3
iv) model evaluation by using fit indices, when three alternative forms of prior knowledge were used as explanatory variables to address research objective 5(b), which is explained in Subsection 3.7.4.

3.7.1 Variance Component Analysis

This subsection provides a rationale for the choice of modelling hierarchical data by a Two-level SEM as represented by equations (3.4) to (3.6) in Subsection 3.3.3 involving variance component analysis of the dependent variables, ACH and SEI (Goldstein, 2011).

Its main purpose was to estimate variances in hierarchical data on ACH and SEI for two components, WITHIN Class and BETWEEN Classes, and examine if the intraclass variation (the proportion of variance accounted for by the BETWEEN Classes component) warranted the use of two-level or higher modelling of students' achievement (10% or more as a rule of thumb) (Goldstein, 2011).

The procedures for variance component analysis are based on the estimation of a simple linear regression equation of the form:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \ldots\ldots. (3.7)$$

For the i_{th} student in the j_{th} class:

Y_{ij} represents, for example, ACH, as measured by postscores (or SEI as measured by expected scores)

β_{0j} represents the class-intercept or the average score of students

ε_{ij} is the residual term

Under the assumption of random intercepts, each class-intercept (β_{0j}) varies around the grand intercept (or the grand mean, β_0) or an average score of all students in the sample. Furthermore, if μ_{0j} is the deviation of the j_{th} class-intercept from the grand mean, then:

$$\beta_{0j} = \beta_0 + \mu_{0j}, \text{ and hence}$$

$$Y_{ij} = \beta_0 + \mu_{0j} + \varepsilon_{ij} \quad \ldots\ldots\ldots. (3.8)$$

As an example, for ACH, Figure 3.4 shows class-intercepts for j=3 and their grand intercept. If we assume that the variation in ACH is only among the intercepts of individual classes (j=1,2,3) and not in their slopes, the regression lines are parallel to each other for three classes. Otherwise, they could have different slopes, as is usually hypothesised in a random-slope model. The grand mean of these three classes could lie anywhere between the class-intercepts, depending on the magnitudes involved (Goldstein, 2011).

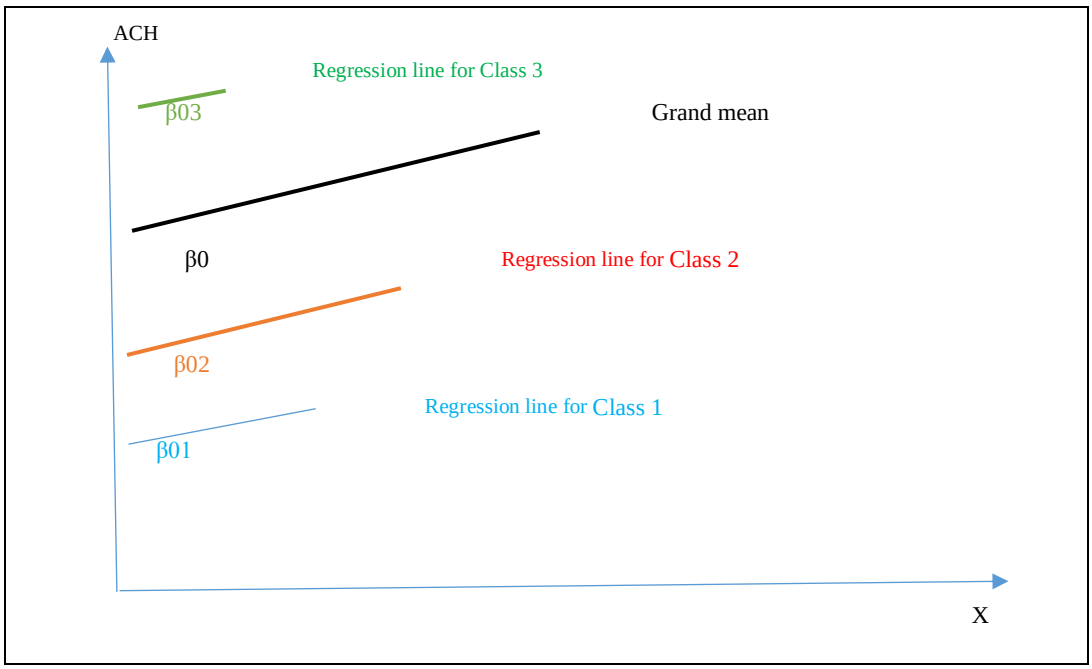


Figure 3.4: A Diagrammatic Representation of Regressions Involving Hierarchical Data of Three Classes and the Grand Mean.

In the present context, regression equation (3.8) was estimated by the maximum likelihood (ML) method to obtain the estimates of variance components:

- the grand mean (β_0)
- the class-intercept (μ_{0j}) and
- the residual (unexplained) variance (ε_{ij}).

In this model, the total variance in each of the dependent variables, ACH and SEI, was decomposed and accounted for variance among three constituents: i) WITHIN Class ii) BETWEEN Classes and iii) residual component, and estimated by using the M*plus* software program by the method of maximum likelihood

For estimating variance components six different sample sets were considered:

1) boys' sample with N=285 and j=14
2) girls' sample with N=179 and j=13
3) overall sample of all students combined with N=464, j=15 classes
4) sample of low-ability students with N=51, j=15 classes
5) sample of medium-ability students with N=110 and j=15 classes
6) sample of high-ability students with N=303 and j=15 classes.

3.7.2 Estimation of Direct and Mediated Effects from Mediation Model

The mediation model, as shown by equations (3.1) to (3.3) in Subsection 3.3.2, is reproduced below with an explanatory variable X (which represents PK or CLE). It is hypothesised that X directly explains variations in students' achievement (ACH) and the mediator (SEI):

$$ACH = a_1 + c\,X + e_1 \quad \ldots\ldots\ldots\ldots \quad (3.1)$$

$$ACH = a_2 + c'\,X + b\,SEI + e_2 \quad \ldots\ldots\ldots\ldots \quad (3.2)$$

$$SEI = a_2 + d\,X + e_2 \quad \ldots\ldots\ldots\ldots \quad (3.3)$$

where

a_1 and e_1 are, respectively, the intercept and the error terms of ACH regression on X

a_2 and e_2 are those of regression of ACH on X mediated through SEI

a_3 and e_3 are those of regression of SEI on X.

An estimate of c gives the direct effect of X on ACH, while that of c' gives the effect of X on ACH adjusted for mediation (MacKinnon, 2007; MacKinnon, Fairchild, & Fritz, 2007; Raudenbush & Sampson, 1999).

According to MacKinnon and Dwyer (1993), the mediated effect can be calculated from the regression estimates of equations (3.1) to (3.3) either by the product method (b.d) or by the difference method (c – c'). Both methods yield the same value of estimates, because their algebraic equivalence was established along with their respective standard errors by MacKinnon et al. (1995).

For testing the statistical significance of effects, it is assumed that there is no covariance between error terms, and that errors are normally distributed with constant variance. I adopted the product method (b.d) to obtain the mediated effects in my study.

To establish statistical mediation, one of the important assumptions underlying the mediation model is that SEI causes change in ACH, but not vice-versa. That is, ACH does not influence SEI. Bandura (1977) also suggested that there could be reciprocal mediation between SEI and ACH, implying that there could be two-way causality between them:

i) SEI to ACH

ii) ACH to SEI.

To verify this, an attempt was made to investigate the extent of reciprocal causation between the mediator variable SEI and ACH. Subsection 3.7.3 below offers details of estimation of Two-level SEM and fit indices information for the model evaluation.

3.7.3 Estimation of Joint Influence from Two-Level SEM

This subsection provides procedures for the estimation of equations (3.4) to (3.6) of the SEM as in Subsection 3.3.3, as well as detailed information on fit indices used for the evaluation of three variants of the SEM.

Estimation of the Two-level SEM was conducted under the assumption of random intercepts, which involves regressions for the WITHIN Class component (Level 1) of the form:

i) ACH on CLE, PK and SEI

ii) SEI on CLE and PK.

At Level 1, variance in each of the dependent variables, ACH and SEI, is explained WITHIN Class and the effects of explanatory variables, PK and CLE are derived using the estimated regression coefficients both in standardised (β) and unstandardised forms (*b*) .

At Level 2, ACH intercept and SEI intercept, the dependent variables of the regressions are assumed to be random, and the underlying explanatory variables vary only BETWEEN Classes, while remaining constant within a class. These Level 2 explanatory variables could include class size, teaching experience, the coefficient of variation in PK (SD per mean), etc.

For estimation of the SEM, three variants were considered using different types of prior knowledge such as PK1, PK2 and PK, where:

i) PK1 is students' prior knowledge that requires integration of ideas and concepts for solving application tasks of Type P1
ii) PK2 is students' prior knowledge required for solving challenging tasks of type P2
iii) PK is students' prior knowledge required to solve all types of tasks combined.

By pairing them with the other explanatory variable, classroom learning environment (CLE), the three regressions of SEM variants are:

i) SEM1: ACH regressed on PK1 and CLE
ii) SEM2: ACH regressed on PK2 and CLE
iii) SEM3: ACH regressed on PK and CLE.

Regression estimates were obtained by the maximum likelihood method, using the M*plus* software (Muthen & Muthen, 2008). The next subsection provides detailed information about the fit statistics that can be used for the model evaluation.

3.7.4 Model Evaluation Using Fit Indices

For evaluation of the SEM variants, regression estimates of equations of Level 1 and Level 2 were compared, separately, using:

- R^2, which explains the variance in the dependent variable due to included variables and was tested for its statistical significance
- LOG L (or ln L), which is negative, and chi-square values of the baseline model

- appropriate fit indices reviewed by Brown (2014) and Kenny (2015).

For deciding suitable selection criteria, Brown (2014) classified fit indices into three types, named, *Absolute Fit Indices*, *Parsimony Correction Indices* and *Comparative Fit Indices*, which are described below.

Among the absolute fit indices, the Standardised Root Mean Square Residual (SRMR) is generally preferred over chi-square (Kenny, 2015) because it is the standardised difference between an observed and predicted correlation. Because SRMR is the average discrepancy between the correlations observed in the input matrix and the correlations predicted by the model, it assesses whether a hypothesis is reasonable. Hu and Bentler (1999) suggest that a value of 0.08 or below indicates reasonably good fit.

Parsimony correction indices incorporate a penalty function for poor model parsimony (i.e., number of freely estimated parameters as expressed by model, df). This group includes the Root Mean Square Error of Approximation (RMSEA) (Steiger & Lind, 1980). It assesses the extent to which a model fits reasonably well in the population. Kenny (2015) provides its computational formula:

$$\text{RMSEA} = [\sqrt{(\chi^2 - df)}] / \sqrt{[df\ (N - 1)]}.$$

N is the sample size and df is the degrees of freedom. "If χ^2 is less than df, then the RMSEA is set to zero. Its penalty for complexity is the ratio of chi-square to df. The measure is positively biased (i.e., tends to be too large) and the amount of bias depends on the smallness of the sample size and df, but primarily the latter" (Kenny, 2015).

MacCallum, Browne and Sugawara (1996, p. 141) used RMSEA values of 0.01, 0.05 and 0.08 to indicate "excellent, good and mediocre fits, respectively". The value could be as small as possible and close to zero, which indicates a perfect fit. They also cautioned that, "in the sample, it might have a value greater than 0.10 for a given model whose population value is 0.05 (which would not be known). There is greater sampling error for models with small df and low N, especially for the former" (MacCallum et al.,1996, p. 141).

Kenny, Kaniskan and McCoach (2015, p. 486) reported that, "when the cut-off values are used to assess the fit of the properly specified models with small *df* and small sample size, the RMSEA too often falsely indicates a poor fitting model. We recommend not computing the RMSEA for small *df* models, especially those with small sample sizes". Furthermore, the authors recommend that, "A confidence interval can be computed for the RMSEA. Ideally, the lower value of the 90% confidence interval includes or is very near zero (or no worse than 0.05) and the upper value is not very large (i.e., less than 0.08)". The width of the confidence interval is very informative about the precision in the estimate of the RMSEA.

For hypothesis testing of the RMSEA, *p of Close Fit* (PCLOSE) provides a one-sided test for the null hypothesis that RMSEA = 0 .05; this is called a *close-fitting model*. Such a model has a small specification error. The alternative one-sided hypothesis is that the RMSEA > 0.05. So, if $p > 0.05$ (i.e., not statistically significant), the fit of the model is 'close' or worse than close fitting.

Comparative fit indices evaluate the user-specified solution in relation to a more-restricted, nested baseline model (Hu & Bentler, 1998), which include the Comparative Fit Index (CFI) (Bentler, 1990) and Tucker–Lewis Index (TLI) (Tucker & Lewis, 1973), which uses a penalty function for adding freely estimated parameters that do not markedly improve the fit of the model.

The comparative fit index, CFI (Bentler, 1990, p. 238) is given by:

CFI = [d (Null Model) - d (Proposed Model)] / d (Null Model), where the non-centrality measure, d = χ^2 - df, and df is the degrees of freedom of the model.

The Tucker-Lewis index (TLI) is given by:

TLI = [χ^2 / df (Null Model) - χ^2 / df (Proposed Model)] / [χ^2 / df (Null Model) – 1], which has "a penalty for adding freely estimated parameters that do not markedly improve the fit of the model" (Tucker & Lewis, 1973, p. 1).

Kenny and McCoach (2003) suggest that RMSEA improves as more variables are added to the model, and that TLI and CFI are relatively stable but tend to decline slightly.

Other comparative fit indices include the Akaike Information Criterion (AIC) (Akaike, 1973, 1974, 1985) and Bayesian Information Criterion (BIC). AIC is defined by AIC (θ) = -2 ln L + 2k. *k* is the number of "independently adjusted parameters required to obtain θ" (Akaike, 1974, p. 719).

Because AIC is a comparative measure of fit, it is meaningful only when two different models are estimated. Lower values indicate a better fit and so the model with the lowest AIC is the best fitting model. As AIC is an incremental measure, when the model parameters are changed, it is the difference in AIC that really matters. The AIC makes the researcher pay a penalty of two for every parameter that is estimated (Akaike, 1974).

BIC, developed by Schwarz (1978), is defined as $-2L_m + m \ln n$, where n is the sample size, L_m is the maximized *log-likelihood* of the model and m is the number of parameters in the model. "The index takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve this particular degree of fit, by imposing a penalty for increasing the number of parameters" (Schwarz, 1978, p. 461).

While fitting models, it is possible to increase the likelihood by adding parameters, but doing so could result in overfitting. Both BIC and AIC attempt to resolve this problem by introducing a penalty term for the number of parameters in the model. The penalty term is larger in BIC than in AIC. BIC increases the penalty as sample size increases. BIC places a high value on parsimony (perhaps too high), whereas the AIC has a penalty of 2 for every parameter estimated (Schwarz, 1978).

3.8 Chapter Summary

This chapter is summarised in three subsections covering the research objectives of this study. The research paradigms, the mediation model and the Two-level SEM as proposed are summarised in Subsection 3.8.1, while research methods

and instruments used for data collection and sample features are summarised in Subsection 3.8.2. Finally, Subsection 3.8.3 outlines data-analysis procedures adopted for accomplishing the research objectives, including the estimation of direct and mediated effects and the joint influence of students' prior knowledge and classroom learning environment on achievement, as well as the model evaluation criteria using fit indices.

3.8.1 Research Paradigms and Research Model

This chapter examined the research methods relevant for estimation of the joint influence of students' prior knowledge and classroom learning environment on their self-efficacy beliefs and achievement in mathematics. Considering the context of and background to the study (Chapter 1) and literature review of relevant studies (Chapter 2) about issue of student diversity and grouping practices used in mixed-ability classrooms and gender equity issues faced in mathematics education in Australia and India, this study addressed five research questions with appropriate research paradigms and methods (see Table 3.1).

Guided by past research on paradigms and ontology (Guba, 1990, 1994; Kuhn, 1962), this study is situated in the constructivist (interpretive) paradigm to address its qualitative objective of how teachers used students' prior knowledge, as well as the positivist paradigm to examine the quantitative objectives that involved:

a) statistical analyses of frequency distributions of prior knowledge and achievement by ability groups and gender
b) analyses of bivariate and partial correlations between achievement and other covariates such as prior knowledge classroom learning environment dimensions, and self-efficacy beliefs in achieving different types of mathematics tasks
c) hypothesis testing of gender differences in prior knowledge and achievement as well as achievement for each ability group
d) first-order and the second-order factor analyses of MCOLES responses

e) rationale for the choice of Two-level modelling when using hierarchical data by variance component analysis of dependent variables, ACH and SEI.

For assessing direct and mediated effects of PK and CLE on ACH, a mediation model involving SEI as a mediator was posited following MacKinnon (2007) and MacKinnon et al. (2007).

Following Bandura (1977, 1980) and Urdan and Schoenfelder (2006), the research hypotheses were considered in my study and a Two-level SEM was adopted to depict the joint influence of the two variables of students' prior knowledge required for learning mathematics (PK) and the psychological construct of classroom learning environment (CLE), which are mediated through students' self-efficacy beliefs in achieving (SEI) on achievement (ACH). Thus, the model has four variables for which four instruments were developed for data collection, as explained in the next subsection.

3.8.2 Instruments for Data Collection and Sample Features

For the first research objective which involved qualitative information, a template of Teachers' reflective Journal (TRJ) was designed for the convenience of teachers who participated in the study to record their classroom teaching experiences in catering for the cognitive needs of their students.

Prior to teaching the selected topic of measurement, teachers used a knowledge framework (Hailikari, Nevgi, & Lindblom-Ylänne, 2007)) (see Appendix 1) to design a pretest (Appendix 2) to assess students' prior knowledge on this topic and to divide students into three ability groups based on prescores. They supplemented classroom teaching by selecting appropriate mathematics tasks. They were supplied with a template of TRJ that has: four questions which required teachers to clarify their learning intentions for their class(es); a brief description of teaching tasks used in the class; instruction about what kinds of questions were raised in the class to encourage participation of low- , medium- and high-ability students; and how teachers

facilitated questioning in the class and encouraged students to ask questions whenever they wanted help in solving tasks or understanding new ideas. .

For measuring students' achievement, a posttest (Appendix 4 for Class 1 and Appendix 6 for Classes 3-16) was designed containing six different types of tasks: four types of tasks as suggested by the knowledge framework (Appendix 1); and two more types for testing students' achievement on more challenging tasks that require higher-order thinking at the year 10 level.

For measuring the psychological construct of classroom learning environment, the new 56-item seven-dimensional MCOLES was developed based on theories of mathematics learning (Dochy, 1992, 1994, 1996; Dochy & Alexander, 1995; Vygotsky, 1962, 1978), constructivist theory (von Glasersfeld, 1985, 2000), Ernest's (1991, 1998) philosophy of social constructivism, and the six principles of teaching and learning adopted by the Department of Education of the State Government of Victoria (2012) and National Curriculum Framework (NCF-2005) implemented by NCERT (2012) in India. The MCOLES was developed by modifying the 88-item 11-dimensional COLES, which is underpinned by constructivist theory and Moos' (1974, 1979) notion of psychosocial environments (Aldridge et al., 2012).

The items of the MCOLES were pilot tested with 20 low-, 20 medium- and 15 high-ability students, which accounted for over 10% of the main sample. To establish its content validity, personal interviews and feedback were obtained from teachers, following Munby (1997). Some students from the Australian school made useful observations and proposed modifications for certain items, which contributed to a better understanding and improvement of the MCOLES.

Students' perceptions of self-efficacy beliefs in achieving different types of given tasks (SEI) were measured by using the MSES, which was developed by following Bandura's (1977, 1980, 1986, 2006) social cognitive theory and self-efficacy theory involving classroom learning environment. Students responded to a set of 31 mathematics tasks, specially designed in the resource material for the MSES for the topic of measurement taught by their teachers at the Year 10 level. The content validity of the MSES was established by using the knowledge framework (Hailikari,

Nevgi, & Lindblom-Ylänne, 2007) and principles of assessment of *equity* and *fairness*, while face validity was satisfied by feedback from experienced teachers and students in a pilot survey.

Students responded to the items (tasks) in the resource material of the MSES on a scale of 1 to 5, where 1 refers to that the task being very difficult (so that the expected score is 0%) and 5 indicates that the given task is very easy or that the student is very sure about solving that task successfully (expected score of 100%). Using this calibration scheme (see Figure 3.2), students' responses to the MSES, which reflect their self-efficacy judgements, were converted into expected scores and used as a measure of their self-efficacy expectancies (SEI) (Bandura et al., 1982; Hackett & Betz, 1989).

The same resource material was used as the posttest, which was administered by teachers to their classes. Teachers evaluated students' solutions to the tasks on the posttest to award them postscores, which were shared with the researcher to use as a measure of achievement when addressing research objectives.

To administer the instruments, a purposive sample of 531 Year 10 mathematics students and their 11 mathematics teachers were selected, based on a rationale explained in Section 1.2, in two stages from one Australian and five Indian secondary schools. At the first stage, the principals of all secondary schools in the Western Regional Victoria, Australia and selected secondary schools in India were sent invitation letters so that those who accepted on a voluntary basis could be recruited. Then, in the second stage, specific letters of invitation were sent to parents, students and teachers for participation in the study, and those teachers and parents of students who returned written consent were included in the sample.

As the study required students to respond to the MCOLES and the MSES, as well as a pretest and a posttest on a topic selected for teaching which was conducted over a period of two to three weeks, the number of students who complied with all requirements varied slightly. Thus, out of 531 students who accepted invitation, only 511 students including 198 girls and 313 boys provided full information for the 56-item MCOLES; 491 students including 184 girls and 307 boys returned responses to

all 31 items for the MSES. However, because a sample size of 464 students including 179 girls and 285 boys complied with all requirements, that complete dataset was used for analyses, including the estimation of the mediation model and Two-level SEM.

The sample, which was drawn during 2015-2016, had Year 10 mathematics students from 15 different classes in six secondary schools, with an overall male-female ratio of 1.6:1. One boarding school in Puttaparthi, a semi-urban town in the State of Andhra Pradesh, India, had 69 students, with boys only in two classes and girls only in another class. All students of this school, being hostellers, also received tutoring support from a dedicated team of resident teachers. Five other schools are co-educational and run by a private society since 1948 in the National Capital Region of Delhi, India.

All students sampled were taught mathematics at Year 10 by both male and female teachers whose teaching experience ranged from 10 to 40 years. The class size varied from 14 to 45 students. The next section briefly presents qualitative and quantitative data-analysis procedures adopted for achieving the objectives of this study.

3.8.3 Estimation Procedures and Model Evaluation

The first research objective of how teachers used students' prior knowledge required for learning a new mathematics topic, being situated in the constructivist (interpretive) paradigm, was addressed qualitatively. It involved a procedure to seek teachers' experiences in teaching a topic of measurement in the mixed-ability classrooms after they assessed students' prior knowledge in a pretest and used the prescores to divide their students into three ability groups. A visual inspection of TRJ notes in a tabulated form was proposed as being suitable for interpreting the reality and basis for examining this research objective.

My research objectives 2(a) to 2(c) of Subsection 1.5.2 relate to gender differences in students' prior knowledge and achievement. Students' prior knowledge (prescores) and achievement (postscores) were collected by using the pretest and posttest (Appendix 4 for Class 1 and Appendix 6 for other classes). The data-analysis

procedure involved statistical frequency distributions by ability groups and gender, and an analysis of correlations (bivariate and partial) between prior knowledge and achievement. The *t*-test (Snedecor & Cochran, 1989) and effect sizes (*d*) (Cohen, 1988) were computed to examine gender differences in prior knowledge and achievement, as well as improvement (or decline) in learning, for the whole sample, individual schools and selected classes, as described in Subsection 3.6.3.

For accomplishing research objectives 3(a) to 3(d), the MCOLES was developed and procedures for establishing translational and criterion-related validities as detailed in Subsection 3.6.4, involved Cronbach's alpha coefficient for establishing construct validity or scale consistency. To retain an item within a scale, a cut-off value of $r > 0.4$ was applied when examining interitem correlations. To establish discriminant validity of a scale, partial correlations of scale means were first computed along with their average value over the remaining six scales, then compared with a cut off value of 0.3. In one-way ANOVA with class membership as independent variable, eta^2 ratio was tested for significance for establishing the concurrent validity, and an analysis of correlations was conducted between students' achievement and different dimensions of classroom learning environment for verifying the predictive validity.

For examining the factorability of the MCOLES, exploratory factor analysis (EFA) was applied to individual dimensions to identify which item combinations yielded high factor loadings, followed by confirmatory factor analysis applied to all dimensions jointly to obtain a solution of factor values. Also, second-order factor analysis was considered to identify a parsimonious factor solution when the first-order factors were found to be highly correlated, theoretically and empirically.

For examining research objectives 4(a) to 4(d) in Subsection 1.5.2, the MSES was developed with the resource material of mathematics tasks, based on sound theories (Hailikari, Nevgi, & Lindblom-Ylänne, 2007) and the twin principles of assessment of learning, *equity* and *fairness*. The validation procedures involved seeking students' and teachers' feedback during a pilot survey. The MSES was structured to the tasks given in the resource material so that students' self-efficacy

judgements as revealed by their responses were calibrated into expected scores. The MSES responses, collected on a scale of 1 to 5 (where 1 refers to that task being very difficult to solve or an expected score of 0%, and 5 refers to an easy task or an expected score of 100%) were calibrated to derive expected scores for the posttest.

For accomplishing research objective 4(b), the procedure to identify underachievers or those whose efficacy-expectancies are lower than achievement was detailed in Subsection 3.6.5. To address research objective 4(c), it was proposed to calculate effect sizes (*d*) for gender differences in self-efficacy expectancies (expected scores) using Cohen's (1988) formula and analyse correlations between i) expected scores and prescores and ii) expected scores and postscores.

For addressing research objective 5(a), the construct of self-efficacy was employed as a mediator, following MacKinnon et al. (1995) and MacKinnon et al. (2007), in the mediation model for estimating the direct and mediated effects of classroom learning environment and students' prior knowledge on achievement.

To address research objective 5(b), a SEM was hypothesised (see Figure 1.1) based on a theoretical rationale outlined in Subsection 1.4.2, following Bandura (1980) and Urdan and Schoenfelder (2006) to estimate the joint influence of students' prior knowledge and classroom learning environment on their achievement.

To quantify the effects using hierarchical data, Goldstein (2011) suggests at least Two-level modelling if the intraclass variation is 10% or more. So, a variance component analysis was proposed for WITHIN Class, and BETWEEN Classes in ACH and SEI variables. To derive the ML estimates of direct and mediated effects from the mediation model, and joint influence from Two-level SEM, the M*plus* software program was proposed to be used.

Three different variants, SEM1- SEM3, were proposed by using different combinations of prior knowledge of solving tasks of Type P1 (PK1), of Type P2 (PK2) and all tasks together (PK) along with CLE at Level 1 (WITHIN Class). For Level 2, variables which remained constant WITHIN Class, but varied BETWEEN Classes, were considered. To evaluate SEM variants, a review of fit indices by Brown (2014)

and Kenny (2015) (Subsection 3.7.4) suggested the use of SRMR, RMSEA, CFI, AIC as appropriate, apart from other statistics (viz., the R^2 and LOG L).

The next chapter deals with results of estimation that address research questions 1 and 2.

Chapter 4

HOW TEACHERS USED THEIR STUDENTS' PRIOR KNOWLEDGE FOR CLASSROOM TEACHING, GENDER DIFFERENCES IN PRIOR KNOWLEDGE AND STUDENTS' ACHIEVEMENT

4.1 Introduction

This chapter provides results for research objectives 1 and 2(a) to 2(c) of Subsection 1.5.2 in five sections as follows. Section 4.2 presents an answer to the first research question of how teachers used their students' prior knowledge of a mathematics topic in classroom teaching, with an example of the Class 1 teacher using prescores for dividing the class into three ability groups. Later, it presents a qualitative analysis of feedback information from all teachers' reflective journals (TRJs) (see Subsections 3.4.2 and 3.6.2).

Section 4.3 deals with research objective 2(a) concerning gender differences in students' prior knowledge and achievement, while Section 4.4 examines research objective 2(b) by an analysis of correlations between students' achievement and i) prior knowledge, ii) teaching experience and iii) class size. Section 4.5 addresses research objective 2(c) about improvement (or decline) in students' achievement using prescores and postscores for i) the whole sample, ii) each school, iii) each class and iv) each ability group in selected classes. Section 4.6 summarises results for the first two research questions.

4.2 Prior Knowledge Assessment and Teachers' Reflective Journal

This section presents how the Class 1 teacher used her students' prior knowledge for the topic of measurement for dividing them into three ability groups in

Subsection 4.2.1. Later, it presents a qualitative analysis of the feedback information from all teachers in Subsection 4.2.2.

4.2.1 Teachers' Use of Prior Knowledge Assessment (Pretest)

As described in Subsection 3.4.1, a pretest was designed by the researcher after consulting teachers for assessing students' prior knowledge required for learning the topic of measurement from the Australian mathematics curriculum (ACARA, 2012, 2015) and the Indian curriculum (NCERT, 2015). As shown in Appendix 2, four types of tasks were selected: D1: Declarative knowledge of concepts and meanings; D2: Declarative knowledge of facts; P1: Procedural knowledge of integration of concepts, meanings and facts; and P2: Procedural knowledge of problem solving applications.

In order that students' prescores would be comparable among all students, a common pretest was designed and administered by teachers, who scored the pretest and used prescores for dividing students into three ability groups:

i) low-ability students who scored less than 40%
ii) medium-ability students who scored 40-60%
iii) high-ability students who scored 60% and above.

This grouping enabled teachers to cater for their students' cognitive needs by selecting suitable supplementary tasks for teaching. For example, Table 4.1 shows how students in Class 1 were divided into three ability groups based on their total prescores. The teacher of this class of 21 students (8 boys and 13 girls) divided students into three groups and provided a suitable teaching method with a lesson plan of appropriate mathematics tasks targeting each group. In turn, this helped the teacher to prepare for teaching the new mathematics topic using a constructivist-orientated approach.

The information about students' prior knowledge can alert teachers when selecting suitable tasks from classroom teaching. They might be able to ensure that the tasks selected are not too easy, repetitive, boring or difficult for any ability group, but

Table 4.1: Prescores (%) for Four Types of Tasks of the Year 10 Topic of Measurement for Class 1

Student Number	Student ID	Types of Tasks (%)* D1	D2	P1	P2	Total scores (%)	Ability Group
Boys							
1	101005	62.5	37.5	37.5	0.0	34.4	Low
2	101013	75.0	50.0	50.0	0.0	43.8	Medium
3	101012	62.5	62.5	62.5	0.0	46.9	Medium
4	101002	75.0	62.5	62.5	0.0	50.0	Medium
5	101018	75.0	62.5	62.5	37.5	59.4	Medium
6	101009	75.0	50.0	50.0	75.0	62.5	High
7	101022	100.0	62.5	62.5	37.5	65.6	High
8	101010	87.5	100.0	100.0	75.0	90.6	High
	Mean	76.6	60.9	60.9	28.1	56.6	
	CV**	0.2	0.3	0.3	1.2	0.3	
Girls							
9	101007	62.5	0.0	0.0	0.0	15.6	Low
10	101019	75.0	0.0	0.0	0.0	18.8	Low
11	101004	87.5	6.3	6.3	0.0	25.0	Low
12	101016	100.0	0.0	0.0	0.0	25.0	Low
13	101020	100.0	12.5	12.5	0.0	31.3	Low
14	101021	75.0	37.5	37.5	0.0	37.5	Low
15	101006	75.0	37.5	37.5	0.0	37.5	Low
16	101011	75.0	37.5	37.5	0.0	37.5	Low
17	101014	62.5	50.0	50.0	0.0	40.6	Medium
18	101017	75.0	50.0	50.0	0.0	43.8	Medium
19	101008	75.0	50.0	50.0	37.5	53.1	Medium
20	101024	100.0	75.0	75.0	37.5	71.9	High
21	101015	75.0	100.0	100.0	37.5	78.1	High
	Mean	81.8	41.5	41.5	10.2	43.8	
	CV**	0.2	0.7	0.7	1.7	0.4	

Note; * D1: Declarative knowledge of concepts and meanings, D2: Declarative knowledge to recall facts, P1: Procedural knowledge of integration of concept meanings and facts, P4: Procedural knowledge of applications of problem-solving tasks.
** CV represents the coefficient of variation in scores (= SD / mean).

interesting for every group, and that no group is left behind in learning. More importantly, teachers could bear in mind the possibility that what is a challenging task to one group could be an easy or known task to a high-ability group, which enables

scaffolding with what students already know about the requisite concepts and meanings to what is ‘higher-order thinking’ for the specific ability group.

For all tasks combined, prior knowledge assessment suggested that boys had a higher mean score (56.6%) than girls (43.8%) (Table 4.1). To analyse this further, consider the types of tasks in the prior knowledge framework for which students displayed a varying degree of performance. For all boys, the mean score was lower at 28.1% for the task type P2, for which they needed problem-solving skills using procedural knowledge. This was in contrast to their relatively better performance (76.6%) on simple tasks of type D1 that required declarative knowledge of concepts and meanings. This result therefore alerted the Class teacher to pay greater attention to problem-solving tasks of type P2 while teaching the new topic by giving those students adequate practice and techniques appropriate for them.

In the case of girls, the mean score was higher for task type D1 (81.8%) than for all other task types, whereas it was the lowest at 10.2% only for the task type P2, showing that most girls were failing to master word problems that involve problem-solving tasks requiring procedural knowledge. Similarly, the low- and medium-ability groups could not attempt any such problem-solving tasks at all. Thus, this finding pointed to the need for the teacher to select tasks of type P1 and P2 that are relatively easier for low- and medium-ability groups than those selected for high-ability group.

Overall, the prescores for Class 1 indicated that all students needed greater attention and support from their teacher for dealing with problem-solving tasks that require procedural knowledge of integrating ideas and skills at their ability levels. A comparison suggests that boys fared better than girls in doing all four types of tasks, although they required greater attention and help when doing problem-solving tasks of type P2.

Similar reports of results for other Classes 3-16 were also prepared and handed to the respective mathematics teachers prior to teaching the topic for the benefit of their classroom teaching during the study. Using their students’ prior knowledge, other class teachers organised their teaching plans by considering types of tasks suitable for their students’ ability levels. Such teachers’ efforts can be an important source of

inspiration for improving their students' understanding of the new lessons taught and can directly impact their achievement on the posttest. Teachers' experiences are analysed in the following subsection using TRJ notes.

4.2.2 Results Based on Teachers' Feedback in TRJ

To address the research objective "to investigate qualitatively teachers' classroom actions that incorporate students' prior knowledge in mathematics", all participating teachers were requested to record their experiences of using students' prior knowledge for making suitable task selection in their individual teaching plans and for addressing the cognitive needs of ability groups, but in a mixed-group setting in their classes. They were also requested to complete the TRJ as shown in Appendix 3 and share their experiences with the researcher at the end of teaching the selected topic. The objective of the TRJ was clearly stated as: *To be able to use students' prior knowledge for teaching a mathematics topic of year 10: Measurement*. It posed four different questions to which all 11 participating teachers gave responses. Five teachers gave responses in writing and six teachers were involved in telephonic personal interviews. Their responses are organised question-wise in Table 4.2.

The first question of the TRJ was about the learning intention for teaching the topic of measurement. The responses by all teachers were uniform and clear across all classes and schools in the sample. For example, the teacher of Class 1 recorded her classroom learning intention as: *We are learning to solve problems involving surface areas & volumes of composite shapes*. This is consistent with the curriculum delivered by mathematics teachers for Year 10.

The second question of the TRJ asked teachers to give a brief description of teaching tasks related to the topic taught in the class. For Class 1, for instance, the teacher described different types of tasks such as: i) area of basic shapes; ii) conversion of units for length, area and volumes; iii) applications of Pythagoras' theorem;

Table 4.2**:** Comments from TRJs on the Use of Prior-Knowledge as a Tool for Classroom Teaching

	Question in the Journal	Class ID	Description of Teachers' Responses
1.	Learning intentions	1	We are learning to solve problems involving surface area & volume of composite shapes
		3-16	In our day-to-day life, we come across various solids which are combinations of two or more solids. Hence, in this chapter, each child should know about the volume and surface areas of the things they use.
2.	Teaching tasks used in class	1	i) Areas of basic shapes; ii) Pythagoras' theorem; iii) Conversion of units including areas and volumes; iv) Strategies for working out surface areas of composite shapes; v) Formulae for the total surface area and volume of cylinders and pyramids etc.
		3-16	Teachers recalled different geometrical shapes by showing physical objects for visualising them; the teaching aids used were: test tubes, conical flask and funnels.

(contd.)

Table 4.2 **(**contd..2**):** Comments from TRJs on the Use of Prior-Knowledge as a Tool in Classroom Teaching

Question in the Journal	Class ID		Description of Teachers' Responses
3 Questions raised for encouraging participation of			
a) Low-ability group	1	i)	Lower level tasks of basic shapes only were used in class and, when they were showing more confidence, easier composite shapes were offered;
		ii)	Scaffolding of multi-step tasks were given (e.g., find missing sides using Pythagoras' theorem, calculate the perimeter of the composite shape);
		iii)	Easy-to-access orientated tasks were given;
		iv)	For reinforcement of ideas learned in the class, suitable homework tasks were given.
	3-16		By way of motivation, students were asked to identify basic shapes of objects such as: i) a tent set up in marriage parties; ii) a medicine capsule consumed by a patient; iii) a water tank used by the Delhi Jal Board for transporting. Later the formulae to work out the surface areas and volume of relevant objects matching with cone, cylinder and hemisphere (bowl, etc.) were recalled. Some teachers adopted a method of storytelling related to the topic because their students could not visualise easily. Afterwards, the basic shapes connected to the story were identified and applied in examination of areas and volumes.

(contd.)

Table 4.2 (contd..3**):** Comments from TRJs on the Use of Prior-Knowledge as a Tool in Classroom Teaching

Question in the Journal	Class ID	Description of Teachers' Responses
b) Medium-ability group	1	'At level' tasks were given with a log sheet to regulate students' own learning; Extra help/scaffolding was given as students requested. Open-ended tasks of a difficult nature, requiring multi-stepped working, were offered.
	3-16	Many students identified the geometric shapes of given objects but struggled in applying the correct formulae for calculating areas and volumes. Then they were asked relevant questions about the dimensions of objects to explain the formulae and make them understand the multi-stepped computation of the areas and volumes.
c) High-ability group	1	Students were given a log sheet of questions to work through from their textbook exercises on Chapter Review and were asked to attempt only those challenging questions for which they needed more help. They were set open-ended tasks and problem-solving tasks involving multi-steps of composite shapes of cones /pyramids etc. Also, they were offered challenging problems that were 'advanced at the year 10 level'.
	3-16	Few students recognised the basic shapes of geometrical objects easily, carried out the tasks on the conversion of units, and calculated areas using the formulae. Then, to encourage higher-order thinking, teachers gave them more challenging tasks involving areas and volumes of composite shapes such as rocket, toys, etc.

(cont.)

Table 4.2 (contd..4): Comments from TRJs on the Use of Prior-Knowledge as a Tool in Classroom Teaching

Question in the Journal	Class ID	Description of teachers' responses
4.How teachers encouraged students to ask questions and facilitated learning	1	Particularly identified those students who were struggling and were offered help. Thus, some students who were shy and did not ask questions publicly in front of the whole class came forward to learn by asking for help. Some students were given modified work if the earlier tasks were found to be unsuitable, either because they were too easy or too difficult for them (e.g. teachers offered life examples like cow paddock, fences, crop areas that suited the farming community background); they generally encouraged them by inspiring (e.g., 'You can do it') or by allowing them to work collaboratively in the classroom where they liked working with the peers of their choice.
	3-16	Teachers used such tasks of practical experiences in day-to-day life to help them improve their learning (e.g. coffee mug, juice glass, jewel box design, and thermometer).

iv) strategies for calculating areas of composite shapes and v) formulae for the total surface area and the volume of cylinders, pyramids, etc. Teachers of other classes also explained their learning intentions by showing various 3D objects from daily life that represented different geometric shapes. These responses described an effective way of communication to a large mixed-ability group of students because they alerted students to the need for learning the formulae for working out the areas of basic shapes. This approach also mirrors the constructivist method of teaching (von Glasersfeld, 1985) because students were required to use their prior knowledge of the formulae of areas and volumes of basic shapes for figuring out the total area and volume of composite shapes/solids by an application of suitable strategies, such as splitting the composite shapes into familiar geometric shapes that they learned at earlier stages.

The third question of the TRJ required teachers to record what questions were raised in the class to encourage participation/discussion from students. For the low-ability group, the teacher of Class 1 made use of basic geometric shapes to introduce applications of Pythagoras' theorem with appropriate scaffolding by requiring students to work out the lengths of missing sides. When more students expressed confidence in solving such tasks with prompts and support from their teacher, they were given open-ended tasks (Sullivan & Mornane, 2014). Their ideas were reinforced with additional practice by giving a choice of homework on similar tasks. For the medium-ability group, the teacher chose slightly more-difficult open-ended tasks after providing an exemplar in class that required multi-stepped procedural knowledge.

For the high-ability group, the teacher used their textbook curriculum and supplied log-sheets of questions from the *Chapter Review* of the topic that included more-challenging and investigative questions requiring higher-order thinking and procedural knowledge. The teacher mentioned that some students of high ability worked on the advanced-level questions at the Year 10 level, which were open-ended problem-solving types of tasks (Sullivan & Mornane, 2014) that could have more than one answer.

In other classes, teachers also used physical objects of known geometric shapes for teaching their low-ability groups of students. For example, they used a tent or medicine capsule and demonstrated the underlying composite geometric shapes of

daily life to motivate and teach them how to calculate surface areas and volumes. For the medium-ability group, teachers found that it was relatively easy for students to identify matching shapes, but difficult to apply appropriate formulae to calculate respective areas of composite shapes. Then teachers gave prompts about the formulae to be used. In the case of high-ability groups, teachers found that students could give suitable strategies to partition the composite shapes into known geometric shapes and could even attempt some problem-solving tasks. Some students were also able to give similar examples of the composite shapes such as rocket and toys or vegetable shapes such as lettuce and cucumber. They also identified the formulae to work out the respective areas and volumes of such shapes. To reinforce their understanding, teachers supplied more-challenging tasks.

The fourth question of the TRJ involved how teachers encouraged students to ask questions, facilitated their learning and requested feedback on any other observations. For this, the Class 1 teacher encouraged each student by targeting him/her on a one-one basis. She met with students who felt too shy to ask for help publicly. For low-ability students, the teacher often offered modified work involving visual learning type of tasks (e.g. working out the perimeter and area of basic shapes with the aid of physical objects familiar in daily life). Because some students came from farming backgrounds, suitable classroom examples included cow paddocks, fences and crop areas. She generally inspired students by saying that “you can do it” or working collaboratively in the classroom.

In other classes, many teachers also encouraged students by using visual learning aids. For example, an ice-cream cone was used to explain how to work out the lateral area of a cone shape, and a cone and bucket were used to explain the shape of frustum. Teachers of some classes mentioned that their students were also given projects involving preparing different geometric shapes and work out their areas and volumes. For example, they used juice glasses, coffee mugs and even card-board to make lunch boxes.

Higher-ability students were also entrusted with designing innovative shapes for working out the perimeter, area and volume of objects encountered in daily life while shopping. For instance, teachers asked students to design a jewellery box and

work out its perimeter. Also, students were asked to find how much cloth would be needed in manufacturing an umbrella of given size. Some teachers observed that, in their classrooms, many students enjoyed and participated in discussions about learning geometric shapes. Students inquired into the relevant formulae for finding areas and volumes when practical relevant experiences from daily life were elicited and involved in their learning.

4.3 Gender Differences in Students' Prior Knowledge and Achievement

My second research question has three objectives 2(a) to 2(c) as mentioned in Subsection 1.5.2. This section addresses research objective 2(a) concerning gender differences in students' prior knowledge and achievement. I examined gender differences in i) students' prior knowledge using prescores and ii) achievement using postscores. For the latter, an assessment of learning (posttest) was developed (see Appendix 4 for Class 1 and Appendix 6 for other classes) which was the same as the resource material for the MSES, as explained in Subsection 3.4.4. Teachers administered the posttest to their students and awarded postscores, which they shared with the researcher. Results of frequency distributions of students' prescores and postscores were analysed and reported for the whole sample and for each school in Table 4.3 with their respective mean and CV values for boys and girls of each school, separately. Results in Table 4.3 reveal that:

i) Mean scores were generally higher for prior knowledge than for achievement, which implies that students did not improve in general (see Figure 4.1). However, there was noteworthy improvement in learning in Schools 1 and 4 for both boys and girls (see Section 4.5).
ii) School 6 exhibited the highest mean scores in prior knowledge amongst all schools surveyed, while girls of School 6 secured the highest mean prescore of 96.20%, which was even higher than for boys (81.70%) at the same school. This performance was followed by girls of School 4 (68.0%) and then girls of School 2 (67.0%).

iii) In the posttest for the overall sample, however, the average performance of boys declined from 67.30% to 59.4%, and that of girls from 66.20% to 55.8%. Boys of School 6, having achieved the highest mean postscore of 81.2% amongst all, continued to perform better than other schools.

Another aspect of comparison was the distributional difference between students within a class, as reflected by the coefficient of variation (CV). The higher the CV ratio (SD / mean), the greater the individual differences and, therefore, the greater the teaching challenge in catering for cognitive needs in a mixed-ability classroom setting and in lesson planning.

Table 4.3: Students' Performance in Pretest and Posttest in Each School

School	Gender	N	Pretest		Posttest	
			Mean (%)	CV	Mean (%)	CV
1	Boys	8	56.60	0.30	62.40	0.35
	Girls	13	39.70	0.47	54.70	0.51
2	Boys	59	71.70	0.20	35.60	0.42
	Girls	36	67.00	0.14	35.80	0.21
3	Boys	56	77.60	0.21	63.50	0.35
	Girls	41	65.60	0.23	66.50	0.24
4	Boys	51	60.00	0.30	71.50	0.28
	Girls	16	68.00	0.31	77.80	0.20
5	Boys	50	42.40	0.27	43.70	0.34
	Girls	39	48.00	0.33	49.30	0.33
6	Boys	61	81.70	0.20	81.20	0.10
	Girls	34	96.20	0.05	61.50	0.17
	All boys	285	67.30	0.31	59.40	0.40
	All girls	179	66.20	0.33	55.80	0.35

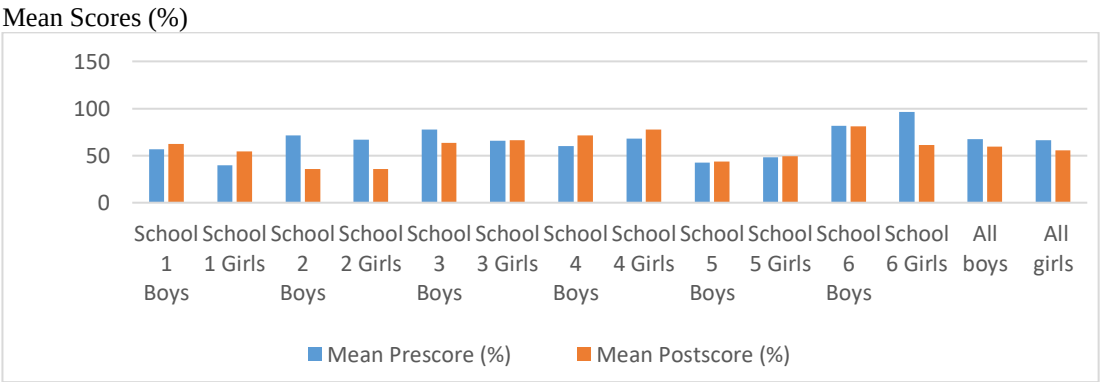


Figure 4.1: Comparison of Mean Prescores and Postscores Separately for Boys and Girls in Each School

For prior knowledge, the CV value for the overall sample was a little less among boys (0.31) than girls (0.33) but, in both cases, appropriate teaching methods were needed for meeting the challenges posed by students of different levels of ability. This also justified the use of ability grouping by teachers in my study based on students' prior knowledge. Despite this effort, for the overall sample, inter-student differences in achievement, reflected by the CV values in postscores, increased more for boys (from 0.31 to 0.40) than for girls (from 0.33 to 0.35).

For examining gender differences (m_1 - m_2), the effect size *(d)* was derived by using the formula (Cohen, 1988):

$$d = (m_1 - m_2) / \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]},$$

where m_1 and m_2 are the mean scores and σ_1 and σ_2 are the standard deviations for boys and girls, respectively. Effect sizes (ES) can be interpreted as how close (or wide) the means are (difference in means) per unit average standard deviation.

Furthermore, a test of statistical significance was conducted for gender differences in prescores and postscores. The relevant procedures for hypothesis testing, as outlined in Subsection 3.6.3, involved *t*-tests.

For the topic of measurement that teachers taught at Year 10, results of gender differences are reported for prior knowledge (prescore) in Table 4.4 and achievement (postscores) in Table 4.5. The first major finding was that there was no significant gender difference in students' prior knowledge required for learning for the overall sample of 464 students (285 boys and 179 girls). Thus, there was nonsignificant gender difference in mean prescores, although the mean score of boys (67.30%) was a little better than that of girls (66.20%). The effect size (*d)* for the gender difference in the mean prescore was only 0.05 standard deviations (SD) for the overall sample. In contrast, the gender difference in achievement was significant because the mean postscore of boys was higher at 59.4% than that of girls (55.8%) on average, with the respective effect size being 0.17 SD (see Table 4.3).

Comparatively, gender differences in prior knowledge were significant for every school except School 4. Whereas boys scored significantly higher than girls in Schools 1-3, girls scored significantly better than boys in Schools 4, 5 and 6. The effect

sizes mirror this finding. For instance, the effect size was positive and highest for School 1 at 0.95 SD, followed by School 3 (0.76 SD) and School 2 (0.39 SD), whereas it was negative for Schools 4 - 6.

Table 4.4: Statistical Significance and Effect Sizes for Gender Differences in Prior Knowledge (Prescores)

School ID	Mean Difference		Std. Error	*t*-value	df (Unequal Var)	Significant at 5%?	Effect Size (*d*)*
	Boys	Girls					
School 1	56.60	39.70	7.9	2.1	16	Yes	0.95
School 2	71.70	67.00	2.4	1.9	92	Yes	0.39
School 3	77.60	65.60	3.2	3.7	90	Yes	0.76
School 4	60.00	68.00	5.8	-1.4	22	No	-0.41
School 5	42.40	48.00	3.0	-1.9	67	Yes	-0.41
School 6	81.70	96.20	2.2	-6.4	77	Yes	-1.20
Overall	67.30	66.20	2.0	0.5	365	No	0.05

*Effect size (*d*) was computed using the formula given by Cohen(1988),
$d = (m_1 - m_2) / \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$

Table 4.5: Statistical Significance and Effect Sizes for Gender Differences in Achievement (Postscores)

School	Mean Difference		Std. Error	*t*-value	Significant at 5%?	Effect Size (*d*)*
	Girls	Boys				
School 1	62.40	54.70	9.7	0.8	No	0.31
School 2	35.60	35.80	1.2	-0.2	No	0.20
School 3	63.50	66.50	3.0	-1.0	No	-0.16
School 4	71.50	77.80	6.7	-0.9	No	-0.35
School 5	43.70	49.30	3.1	-1.8	Yes	-0.36
School 6	81.20	61.50	2.1	9.2	Yes	2.11
Overall	59.40	55.80	1.8	2.0	Yes	0.17

* Effect size (*d*) was computed using the formula given by Cohen (1988).
$d = (m_1 - m_2) / \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$

In School 6, girls had the highest mean score in the prior knowledge test (96.1%) and the lowest inter-student differences as reflected by the CV value of 0.05. Similarly, boys at the same school had the highest mean score of 81.70% in prior knowledge. The effect size for gender differences in prior knowledge for School 6 was very large with girls scoring 1.20 SD higher than boys. It can also be interpreted that, in the prior knowledge test, the mean score of girls was at 88th percentile of the mean score of boys in this school. These findings for prior knowledge were not surprising

for School 6 because most students were disciplined learners, given the school background characteristics as described in Subsection 3.5.2.

Results in Table 4.5 for Schools 1-4 support that gender differences in postscores were statistically nonsignificant although, for School 1, the mean score for boys was noticeably higher at 62.40% than for girls at 54.70%, with an effect size of0.31SD. For School 4, the girls' mean was higher at 77.80% than that of boys at 71.50%, with an effect size of 0.35 SD. Gender differences were significant only for Schools 5 and 6, with girls performing better than boys (effect size of 0.36 SD) in School 5, but boys performing significantly better than girls in School 6 (very large effect size of 2.11 SD), despite the fact that the mean score of girls in this school was the highest in the pretest (96.2%) amongst all students in the sample.

4.4 Associations of Students' Achievement with Prior Knowledge, Teaching Experience and Class Size

This section addresses research objective 2(b) in Subsection 1.5.2 involving the extent to which students' achievement is associated with their prior knowledge, teaching experience and class size for the whole sample as well as for individual classes. Students' prescores and postscores were used to obtain bivariate correlations in Subsection 4.4.1 for the whole sample, and in Subsection 4.4.2 for individual classes by types of tasks on the pretest.

4.4.1 Correlation Analysis of Students' Achievement: Whole Sample

This subsection presents an analysis of correlations between students' achievement and i) prior knowledge, ii) teaching experience, iii) class size and iii) the interaction between teaching experience and class size (product variable). To examine associations of students' achievement with prior knowledge, bivariate correlations between students postscores and prescores were estimated using Pearson correlation coefficients. They were found to be positive and moderate at around 0.39,

with the class mean as a unit of analysis, and at 0.35, with the student as a unit of analysis (Table 4.6).

The correlation between number of teaching years and achievement (postscores) was low at 0.03, with the student as a unit of analysis, but markedly higher at 0.31, with the class as a unit of analysis. This finding is meaningful given that teaching experience remained constant, or the same teacher dealt with all students

Table 4.6: Correlations between Students' Achievement and Other Variables: Class Size, Teaching Experience and Prior Knowledge

Variables	Class as Unit (N=15)	Student as Unit (N=464)
Teaching Years	0.31	0.03
Class Size	0.37	0.20
Interaction between Class size and Teaching years (size x teaching years)	0.51	0.17
Prior Knowledge	0.39	0.35

within a class and that teaching experience varied between classes from a minimum of 10 years to as high as 35 years.

When the class mean was used as a unit of analysis, the correlation between class size and students' achievement was positive and moderate at 0.37 and, when the student was used as a unit of analysis, it was smaller at 0.20. The finding that students' achievement was positively correlated with class size appears contradictory because, with increasing class size, the average time that a teacher can spare to help a student on a one-on-one basis is reduced, given that the class time is limited to a standard lesson period of 45 to 50 minutes. Although the teacher might wish to give as much support as possible to help each student, it might not happen because of limited class time and large student numbers. These results are discussed in detail with other concomitant factors in Chapter 8. Associations between students' achievement and their prior knowledge, when classified by types of tasks, are considered below.

4.4.2 Correlation Analysis of Achievement with Prior Knowledge by Types of Tasks

In this subsection, correlations between students' prior knowledge and their achievement (postscores) were estimated for different types of tasks on the pretest in individual classes of each school, with student as the unit of analysis. Students' prior knowledge was classified according to the four types of tasks D1, D2, P1 and P2 in the knowledge framework (Hailikari, Nevgi, & Lindblom-Ylänne, 2007) (see Appendices 1 and 2), and measured by the respective prescores. Using these prescores and postscores, bivariate correlations between students' achievement and prior knowledge were estimated for each type of task and each class, as reported in Table 4.7.

Another analysis involved students' performance in classes where correlations were not statistically significant, such as Classes 5 and 6. In some classes, the

Table 4.7**:** Correlations between Students' Achievement and Prior Knowledge of Different Task Types in Each School and Class

School ID	Class ID	N	Correlations with Achievement for Types of Tasks				
			D1	D2	P1	P2	All Types***
1	1	21	0.54 *	0.43*	0.43*	0.46*	0.56*
2	3	38	- 0.11	0.38*	0.01	0.36*	0.25**
	4	41	0.26**	0.25**	0.40*	0.31*	0.38*
	5	16	- 0.40**	0.20	0.07	0.40**	0.22
3	6	41	0.08	0.19	0.11	0.18	0.19
	7	17	0.23	0.46*	0.39**	0.11	0.40**
	8	39	0.00	0.20	0.21**	0.00	0.19
4	9	30	0.15	0.03	0.65*	0.71*	0.75*
	10	37	0.61*	0.47*	0.37*	0.22**	0.54*
5	11	34	0.12	0.48*	0.38*	0.67*	0.54*
	12	27	0.18	0.52*	0.37*	0.37*	0.60*
	13	28	- 0.05	0.28	0.31**	0.12	0.28**
6	14	31	0.14	0.04	0.31**	0.35	0.40
	15	30	0.00	0.32**	0.29**	0.09	0.27**
	16	34	0.00	0.00	0.16	0.29**	0.30**
All		464	0.23*	0.24*	0.32*	0.30*	0.35*

* Statistically significant at the 5% level; ** Statistically significant at the 10% level

*** Description of task types, D1: Declarative knowledge of concepts and meanings, D2: Declarativ knowledge to recall facts, P1: Procedural knowledge of integration of concepts meanings and facts, P2: Procedural knowledge of applications of problem solving tasks.

correlations were zero (e.g. Classes 8, 15 and 16) between prior knowledge of type D1 and students' achievement, often because there was no variation in either prescores or postscores. For example, almost all students of Classes 8, 15 and 16 obtained the identical score of 100% in the pretest involving easy items for task D1. Also, the correlation was zero in some classes where almost all students failed to answer any problem-solving tasks of type P2 (e.g., a zero score for Task type P2 was obtained by students in Class 8). Similarly, correlations were negative for Classes 3, 5 and 13, where students scored very low postscores despite having obtained high prescores in the pretest.

4.5 Improvement in Students' Achievement

Research objective 2(c) in Subsection 1.5.2 was to investigate changes in students' achievement relative to their prior knowledge for i) the whole sample, ii) each school, iii) each class and iv) each ability group in selected classes. The changes were analysed mainly to discern differences in achievement patterns. For example, an improvement in achievement noticed in a class could be concealed when achievement is examined by combining all classes of a school, or a decline observed in achievement for a school might not be identifiable when achievement is analysed for the overall sample. Results are presented for each ability group for the whole sample of 464 students in Subsection 4.5.1, for each school in Subsection 4.5.2, for each class in Subsection 4.5.3, and for all ability groups within selected classes in Subsection 4.5.4.

4.5.1 Improvement in Students' Achievement: Whole Sample

This subsection examines whether teachers' selected tasks, suited to their students' ability levels, were useful for improving achievement (postscores) in mixed-ability classrooms. This was inferred from Table 4.8 by an analysis of frequency distributions of the prescores and postscores of students of each ability group, as well as for boys and girls separately, for the selected topic of measurement.

The first finding was that, when each student's score was compared between the pretest and posttest, only 31.7 % or 147 out of 464 students achieved higher scores in the posttest than for the pretest. This was comprised of 60.8% or 31 out of 51 students of low ability; 52.7% or 58 out of 110 students of medium ability; and 19.1% or 58 out of 303 students of high-ability. The range of their improvement was wider (2- 60 %) for the low- and medium-ability groups than the high-ability group (1-36%).

In the high-ability group, Table 4.8 reveals that, although 165 out of 303 students scored between 60-100% in the posttest, only 58 students achieved higher than their pretest score, and hence only about 19% of them improved learning. On the other hand, 54 students obtained less than 40% and 84 students scored between 40-60%. Among the 165 students who obtained between 60-100% that on the posttest, 116 students scored less than on the pretest. Thus, in the high-ability group, for as many as 245 out of 303 students (80.8 %), achievement scores were no better than prior knowledge scores, and hence there was no improvement in learning.

Furthermore, improvement in learning occurred more for boys than girls: 103 out of 285 boys (36.4 %) and 58 out of 179 girls (32.4 %) improved their learning (Table 4.8).

A similar result emerged when students' mean prescores and postscores were analysed for each ability group. Students' performance was better, on average, for the pretest than for the posttest for both boys and girls, separately. The mean prescore and postscore were, respectively, 67.3 % and 59.4 % for boys and 66.2% and 55.8% for girls (Table 4.9). For the overall sample, results at the bottom of Table 4.9 suggest that achievement declined significantly at the 5% level for both boys and girls, with effect sizes being slightly higher for girls (0.50 SD) than for boys (0.35 SD), which suggests that students possessed an acceptable level of prior knowledge which could be scaffolded to enhance a higher level of students' achievement by a constructivist teaching method.

Another important distributional feature was the variances in students' prior knowledge and achievement scores. The prescores varied with a lesser coefficient of variation (CV) for the pretest (0.32) than that for the posttest (0.39) (not shown in

Table 4.8: Students' Improvement in Learning for Each Ability Group, N=464

Ability Group	Number of Students With Postscores			Sample Size	Number of Students Who Improved		Range of Improvement
	0-40%	40-60%	60-100%	N	%		
Low (0 - 40%)	20	18	13	51	31	60.8	6 - 60 %
Medium (40 - 60%)	35	35	40	110	58	52.7	2 - 52 %
High (60 - 100%)	54	84	165	303	58	19.1	1 - 36 %
All Groups	109	137	218	464	147	31.7	1 - 60 %
Boys	70	70	145	285	103	36.1	1 - 60 %
Girls	39	58	82	179	58	32.4	1 - 58 %

Table 4.8). In a mixed-ability setting, the higher the CV, the greater the teaching challenges because teaching must address variations in ability levels among learners. In absolute terms, the variance in learning was comparatively high, as revealed by respective values of CV. Evidently, because of a higher CV for postscores, teaching challenges to cater for the individual cognitive needs of students were more accentuated.

4.5.2 Improvement in Students' Achievement: Each School

For answering research question 2(c) in Subsection 1.5.2, further analysis of students' achievement in each school are reported in this subsection. An important question addressed was: At which school did teachers contribute to a significant improvement in students' learning, given that they taught them in ability groups based on students' prior knowledge?

Table 4.9 shows results for the statistical significance of, and effect sizes (Cohen, 1988) for, differences between prescores and postscores separately for boys and girls in each school. A positive sign for an effect size indicates an improvement, whereas a negative sign reflects a decline in learning. First, in School 6, boys achieved the highest mean score of 81.2% in the posttest, but there was a nonsignificant decline in learning, with an almost negligible effect size of -0.04 SD. Another major finding was that there was a significant improvement in learning among boys in School 4, with

Table 4.9: Effect Sizes and Statistical Significance for Differences between Pretest and Achievement Posttest: Boys and Girls

School	Sex	Mean Scores		Std. error	*t* -value	Effect sizes (*d*)*	Significant?
		Pretest	Posttest				
School 1	Boys	56.60	62.40	8.9	0.6	0.48	No
	Girls	39.70	54.70	8.8	1.7	1.14	Yes
School 2	Boys	71.70	35.60	2.1	-17.3	- 3.56	Yes
	Girls	67.00	35.80	1.8	-17.6	- 4.70	Yes
School 3	Boys	77.60	63.50	2.8	-5.0	- 1.22	Yes
	Girls	65.60	66.50	3.4	0.3	0.90	No
School 4	Boys	60.00	71.50	3.9	2.9	0.90	Yes
	Girls	68.00	77.80	8.0	1.2	0.66	No
School 5	Boys	42.40	43.70	2.3	0.6	0.16	No
	Girls	48.00	49.30	3.6	0.4	0.12	No
School 6	Boys	81.70	81.20	2.9	-0.2	- 0.04	No
	Girls	96.20	61.50	1.0	-35.4	-10.20	Yes
Overall Sample	Boys	67.30	59.40	1.6	-4.8	- 0.35	Yes
	Girls	66.20	55.80	2.1	-4.9	- 0.50	Yes

* Effect size (*d*) was computed using the formula given by Cohen (1988), $d = (m_1 - m_2) / \sqrt{[(\sigma_1^2 + \sigma^2) / 2]}$

the mean score increasing from 60% for the pretest to 71.5% in the posttest, with a large effect size of 0.90 SD. For boys in School 1 also, the mean score increased markedly from 56.6% in the pretest to 62.4% in the posttest.

4.5.3 Improvement in Students' Achievement: Each Class

This subsection considers results of changes in students' achievement in each class in order to address research question 2(c) of Subsection 1.5.2. To compare classes, mean prescores and postscores and their differences (postscore mean – prescore mean), effect sizes and *t*-statistic were used as explained in Subsection 3.6.3 (Cohen, 1988; Snedecor & Cochran, 1989) and reported in Table 4.10. Out of 15 classes, as many as seven classes showed some improvement in learning with the highest effect size (*d*) of 3.1 SD for Class 10 (Table 4.10), followed by Class 1 (0.7SD) and Class 11 (0.3 SD). A major improvement occurred in learning among students of Class 10, as their class average score increased from 54.7% in the pretest to 81.7% in the posttest.

Unfortunately, seven other classes registered a significant decline in achievement with their effect sizes varying between -0.7 SD for Class 6 to -6.0 SD for Class 16. As observed in Subsection 4.4.2, Class 16 of School 6 showed a steep fall in their mean score from 96.2% in the pretest to 61.5% in the posttest, despite the lowest CV value in pretest scores. Also, the mean postscore was less than 40% in Classes 3, 5 and 7, indicating that these classes had many students of low ability, but a review of individual performance of students within these classes indicated no improvement in postscores when compared with prescores.

Another important finding was that the pattern of improvement (or decline) in different classes occurred regardless of teaching experience and class size. All 11 teachers in the sample had taught for 10 to 40 years. For instance, 37 students in Class 10, and 21 students in Class 1 were taught by relatively 'less' experienced teachers who had 13 and 10 years of mathematics teaching service, respectively. Yet, Class 10 exhibited a significant improvement in learning. Similarly, students of Class 1 also achieved a higher mean score of 57.6% on the posttest as comparted to 46.1% on the pretest,

which can be attributed partly to teachers' efforts, despite a large class size, as explained below.

Table 4.10: Class Means, Standard Deviations and Pretest-Posttest Changes (Effect Sizes and Significance) for Each Class in the Sample

Class ID	Class size	Mean score		SD		Change	
	N	Pretest m_1	Posttest m_2	Pretest σ_1	Posttest σ_2	Effect Size (*d*)	Significant at 5 %? (@)
Classes with Nonnegative Effect Sizes Showing Some Improvement							
10	37	51.1	83.9	15.3	14.3	3.1	Yes
1	21	46.1	57.6	19.8	25.3	0.7	No
11	34	52.8	55.6	13.2	12.8	0.3	No
13	28	44.6	45.7	9.8	13.7	0.1	No
14	31	77.9	78.7	14.8	3.9	0.1	No
8	39	77.7	77.9	16.3	14.8	0.0	No
12	27	35.1	34.7	11.6	13.9	0.0	No
Classes with Negative Effect Sizes Showing a Decline							
15	30	85.6	83.7	16.3	10.0	-0.2	No
6	41	70.6	63.1	16.9	15.1	-0.7	Yes
9	30	74.8	59.5	14.2	15.5	-1.5	Yes
7	17	64.9	38.6	15.6	13.1	-2.6	Yes
4	41	75.6	41.8	12.1	13.8	-3.7	Yes
5	16	58.2	31.4	8.7	11.3	-3.8	Yes
3	38	68.7	30.8	11.7	8.6	-5.2	Yes
16	34	96.2	61.5	4.8	10.5	-6.0	Yes
Overall	464	66.4	58.0	21.2	22.6	-0.5	Yes

Notes: @ Using the *t*-test for mean difference (Snedecor & Cochran, 1989)
* Effect size, $d = (m_2 - m_1) / \sqrt{[(\sigma_1^2 + \sigma_2^2) / 2]}$ (Cohen, 1988)

Teaching experience could moderate the effect of class size on students' achievement, given that teachers adopted ability grouping as a strategy in this study. To measure the moderated effect of interaction of teachers with students in a large class, a new variable, *size* x *teaching years*, was derived as shown in Table 4.6. Its correlation with students' achievement was much higher at 0.51 when the class mean was the unit of analysis. Thus, teaching experience works to some extent as a covariate with class size in influencing achievement.

Interestingly, students' achievement declined in smaller classes taught by highly-experienced teachers. For example, in Class 5, students' achievement declined significantly with the effect size being -3.8 SD, despite a small class size of 16 and a low standard deviation of 8.7 in prescores.

Thus, students having prior knowledge for learning a new topic, though necessary, does not ensure higher achievement. This is substantiated by the performance of students in Classes 5 and 7, which were taught by highly-experienced teachers. For Class 5, with a smaller class size of 16 students only, students' achievement was, on average, lower at 31.4 % even though students' had somewhat better prior knowledge as evidenced by a prescore average of 58.2%. Similarly, in Class 7, students' achievement declined from an average prescore of 64.9 % to an average post score of 38.6 % with an effect size of -2.6 SD, despite having a small class size of 17 students only.

4.5.4 Improvement in Students' Achievement for Ability Groups within Selected Classes

To further address research objective 2(c) of Subsection 1.5.2, students' achievement results are elaborated in this subsection for all three ability groups in selected classes that displayed contrasting patterns or improvement and decline. For this, use was made of frequency distributions of test scores, means and CVs (where CV is defined by SD per unit mean of scores within a class and reflects teaching effort in mixed-ability classroom) as displayed in Figure 4.2.

Other things remaining constant, the higher the CV in students' prior knowledge required for learning, the greater is the teaching challenge to deliver the curriculum and, so, a greater effort might have to be invested by the respective class teacher to improve students' learning. Thus, CV reflects the extent of teaching challenge faced by a teacher to address students' cognitive needs within a class.

Results in Figure 4.2 support that teachers faced relatively high teaching challenges in Classes 1, 10 and 12 because their CV of prescores ranged between 0.30 and 0.43, whereas the teacher of Class 16 had a relatively smaller teaching challenge,

because this class displayed a low value of CV of prescores (0.05), which implies that students had more or less the same prior knowledge. For Classes 3, 4, 5, 9, 14 and 15, the extent of teaching challenge faced was moderate, as their CV values ranged between 0.10-0.20. Students' achievement also varied in these classes, as detailed below.

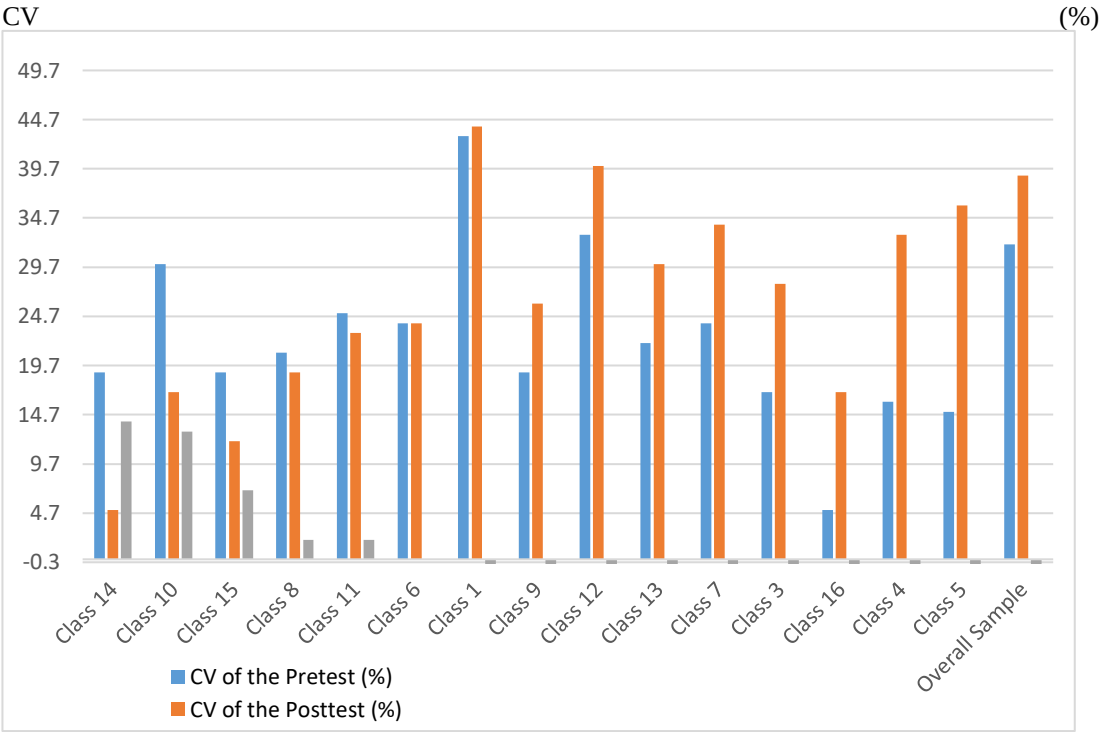


Figure 4.2: Comparison of CV Values between the Posttest and Pretest for Different Classes

Table 4.11 reports a pattern of decline in postscores for Classes 3 and 7. As discussed in Subsection 4.5.3, in both Classes 3 and 7, the mean postscores declined, more in Class 3 to 32.0% from 68.7% than in Class 7 to 38.6% from 64.9 % in the pretest. These two classes had different class sizes; Class 7 had 17 students only and Class 3 had 38 students. Moreover, both classes were taught by very-experienced teachers. The main finding was that students' achievement declined regardless of class size or teaching experience.

In Class 7, there were six students in the medium-ability group and 11 students in the high-ability group. The mean postscores of these 17 students declined with eight students scoring less than 40% and the remaining nine students scoring between 40-60% only. Also, the gap in performance widened between them, as the respective CV value increased to 0.34 in the posttest from 0.24 in the pretest.

In contrast, Table 4.12 presents a pattern of improvement in Classes 1 and 10 which had relatively larger numbers of students (21 students in Class 1 and 37 in Class 10). Yet, in Class 1, the mean postscores increased to 57.6% from a low of 46.1%, despite a high CV value of 0.43 in the prescore. The disadvantage of a larger class size was internalised by these two teachers as they divided their students into smaller groups based on prescores, which probably helped the teachers to focus on teaching students according to their levels of ability and paid off because students showed remarkable improvement in their achievement.

In Class 1, there were nine students in the low-ability group. Possibly those students benefitted from the learning opportunities offered by their mathematics teacher because they achieved higher postscores. Thus, five out of nine students from the low-ability group and five out of seven from the medium-ability group moved into the higher range of 60-100 % in the posttest. Consequently, the mean score of this class improved significantly to 57.6%. However, six out of 21 students still needed adequate teacher help.

In Class 10, most students achieved higher scores of 60-100% in the posttest, which comprised 6 out of 7 students in the low-ability group and 15 out of 18 in the medium-ability group. Also, all 12 students of the high-ability group continued to score between 60-100%. Thus, all 37 students of Class 10 contributed to a steep increase in their mean postscore to 83.9 % from an average prescore of 51.1 %. As a result, individual differences were also reduced as reflected by a decrease in the CV value of 0.30 in prescores to 0.17 in postscores. The following section provides a summary of results for research questions 1 and 2 of this study.

4.6. Chapter Summary

In this chapter, I presented results for research objectives 1, 2(a) to 2(c) of Subsection 1.5.2. Below I summarise findings and limitations in five subsections:

Table 4.11**:** Pattern of Decline in Students' Achievement for Classes 3 and 7

Ability Groups Based on Pretest Scores	Number of Students whose Achievement Varied between				Pretest		Posttest		Effect Size
	0-40 %	40-60 %	60-100 %	N	Mean %	CV	Mean %	CV	*d*
Scenario of Decline in Class 3	30	8	0	38	68.7	0.17	32.0	0.28	-5.20
Low-ability (0-40%)	0	0	0	0					
Medium-ability (40-60%)	5	3	0	8					
High-ability (60-100%)	25	5	0	30					
Scenario of Decline in Class 7	8	9	0	17	64.9	0.24	38.6	0.34	-2.58
Low-ability (0-40%)	0	0	0	0					
Medium-ability (40-60%)	4	2	0	6					
High-ability (60-100%)	4	7	0	11					

Table 4.12**:** Pattern of Improvement in Students' Achievement for Classes 1 and 10

Ability Groups Based on Pretest Scores	Number of Students whose Achievement Varied between				Pretest		Posttest		Effect Size
	0-40 %	40-60 %	60-100 %	N	Mean %	CV	Mean %	CV	*d*
Scenario of Improvement Class 1	6	3	12	21	46.1	0.43	57.6	0.44	0.71
Low-ability (0-40%)	4	1	4	9					
Medium-ability (40-60%)	2	1	4	7					
High-ability (60-100%)	0	1	4	5					
Scenario of Improvement Class 10	0	4	33	37	51.1	0.30	83.9	0.17	3.13
Low-ability (0-40%)	0	1	6	7					
Medium-ability (40-60%)	0	3	15	18					
High-ability (60-100%)	0	0	12	12					

qualitative analysis of how teachers used students' prior knowledge for classroom teaching based on their feedback in TRJ in Subsection 4.6.1; quantitative assessment of gender differences in students' achievement and prior knowledge in Subsection 4.6.2; correlations between students' prior knowledge and achievement in Subsection 4.6.3; changes in students' achievement in Subsection 4.6.4; and limitations in Subsection 4.6.5.

4.6.1 Teachers' Use of Their Students' Prior Knowledge for Classroom Teaching

To examine the first research question, I consulted the 11 participating mathematics teachers of Year 10, selected the mathematics topic of measurement and designed a common pretest of students' prior knowledge in that topic. Teachers administered the pretest and scored it to provide prescores. Then, teachers divided their students into three ability groups, taught the selected topic of measurement over a period of two to three weeks using supplementary tasks suited to students' levels of ability, and recorded their experiences when using students' prior knowledge using the template of TRJ that has four questions.

In response to the first question about the learning intentions for the new topic, all teachers replied saying that they adopted the method of whole-class teaching for introducing the learning intention of the topic of measurement by showing various 3D objects from daily life that represented different geometrical shapes. These responses identified an effective way of communication to a large mixed-ability group of students while motivating them for learning the relevant formulae of perimeter, total surface area and volume of basic shapes. This finding was also consistent with constructivist teaching in which teachers verify their students' prior knowledge before starting to teach any new topic.

The second question of the TRJ was about how teachers taught students in the pre-classified ability groups. Many teachers reported that they gave prompts to motivate low-ability students, as well as assigning open-ended tasks that were challenging and reinforced their understanding with more practice in the classroom and homework on similar tasks.

The third question of the TRJ identified how teachers encouraged students of each ability group to facilitate learning. Teachers replied that they met 1-1 with students who felt too shy to ask for help publicly. For the low-ability group, many teachers often offered modified work by giving visual tasks such as working out the perimeter and area of basic shapes with the aid of physical objects. Teachers generally encouraged students with words like '*you can do it'*, or by working collaboratively with their peers in the classroom. They used visual aids (viz., an ice-cream cone to explain how to work out the lateral area, a cone and a bucket to explain the shape of a frustum) and they assigned projects to students for preparing different geometrical shapes to calculate areas and volumes. In summary, teachers' feedback suggested that many teachers made effective use of their students' prior knowledge in introducing the new topic and taught by a suitable selection of tasks to cater for the cognitive needs of each ability group within their classrooms.

4.6.2 Gender Differences in Students' Prior Knowledge and Achievement

To address research question 2 (a), gender differences in students' pretest and posttest scores were explored using effect sizes and *t*-tests. The first major finding was that, for the topic taught at Year 10 and assessed by teachers, gender difference in students' prior knowledge required for learning was not statistically significant at the 5% level for the overall sample of 464 students (285 boys and 179 girls), although the average score of boys was a little better at 67.30% than that of girls at 66.20%. The effect size (*d)* for gender differences in prescores was only 0.05 SD for the overall sample. In contrast, the gender differences in achievement was significant because the average postscore of boys was higher at 59.4% than that of girls (55.8%) on average, with an effect size of 0.17 SD.

The second finding was that there were wide gaps in prior knowledge, as reflected by the CV ratio of 0.31 for boys and 0.33 for girls, which suggests considerable teaching challenges to cater for the cognitive needs of students of all levels of ability in mixed-ability classroom setting. This also justified the use of ability grouping based on students' prior knowledge for classroom teaching. In School 6,

however, which adopted the path of disciplined learning, the CV value was noticeably small at 0.05 for girls and 0.20 for boys.

The third finding relates to individual schools in the sample. In Schools 1-3, gender differences in prior knowledge were significantly higher for boys than girls. The effect size was positive and the highest for School 1 at 0.95 SD, followed by School 3 (0.76 SD) and School 2 (0.39 SD), whereas in Schools 4 - 6, girls scored significantly better than boys.

For students' achievement in Schools 1- 4, gender differences were statistically nonsignificant although, for School 4, the girls' mean was higher at 77.50% than that of boys at 71.50%, with an effect size of 0.35 SD but gender differences were significant only for Schools 5 and 6, with girls performing better than boys in School 5 (effect size of -0.36 SD), but boys performing significantly better than girls in School 6, with a very large effect size of 2.11 SD.

4.6.3 Correlations of Students' Achievement with Prior Knowledge, Teaching Experience and Class Size

For accomplishing research objective 2 (b) of Subsection 1.5.2, correlations were estimated and found to be relatively high at 0.35 for all types of prior knowledge combined. However, correlations with achievement varied from a low of 0.23 for Task type D1: Declarative knowledge of concepts and meanings, to as high as 0.32 for Task type P1: Procedural knowledge of integration of concepts, meanings and facts. The variation in correlation between achievement and prior knowledge of different tasks seems meaningful and consistent with the fact that prior knowledge of solving application tasks (Type P1) assumes greater importance than other types in learning a new topic, and therefore teaching must emphasise this aspect for improving students' future performance.

The correlation between teaching experience (measured by working years) and achievement (postscores) was also moderate at 0.31, with the 11 teachers having experience from 10 to 32 years.

The correlation between class size and students' achievement was also positive and moderate at 0.37, when the class mean was used as a unit of analysis, and was 0.20 when the student was used as a unit of analysis. Other things remaining constant, students' achievement should be better with smaller class size because teachers could pay greater attention on a one-one basis. However, for the present sample, the positive correlation between class size and students' achievement is plausible, given that teachers adopted within-class ability grouping for teaching and paid individual attention to students in small groups to improve their achievement even in large classes. This was also supported by a moderate effect obtained by taking the teaching experience jointly with class size. The correlation between achievement and the new variable of *class size* x *teaching years* was 0.51.

Moreover, students' achievement was comparatively lower even with a smaller class size and a highly-experienced teacher, as in Classes 6 and 7 with class sizes of 41 and 17, respectively. Surprisingly, students' mean postscores fell in both these classes, from 70.6% to 63.1% in Class 6, and from 64.9% to 38.6% for Class 7 which had a smaller class size of 17 students only.

4.6.4 Improvement in Students' Achievement

This subsection summarises to what extent teachers' selection of tasks suited their students' levels of ability and was useful for improving achievement (postscores) in mixed-ability classrooms. This was inferred by an analysis of frequency distributions of prescores and postscores for students of each ability group.

The first finding was that, when each student's score was compared between the pretest and posttest for examining any improvement in learning, only 31.7 % or 147 out of 464 students achieved higher scores in the posttest than the pretest. This was comprised of 60.8% or 31 out of 51 students of low ability; 52.7% or 58 out of 110 students of medium ability; and 19.1% or 58 out of 303 students of high-ability. This finding lends support to the teaching strategy adopted by participating teachers in my study that involved a judicious selection of mathematics tasks suitable for different ability groups of students based on prior knowledge. Their teaching seems to have catered for the cognitive needs of students to a certain extent, but favoured students of

low and medium ability more than students of higher ability. In the high-ability group, however, a large proportion of 80.8% of students (245 out of 303) experienced a decline in learning relative to their prior knowledge, which needs further investigation.

The prescores had a lower coefficient of variation (CV) of 0.32 than that for the posttest (0.39). In a mixed-ability setting, the higher the CV, the greater the teaching challenges because teaching must address variations in the ability levels of learners. In absolute terms, the variance in learning was comparatively high, as revealed by respective values of CV. Evidently, because of an increase in CV values of postscores, teaching challenges were more accentuated.

By examining performance in individual schools, I found that boys in School 6 achieved the highest mean score of 81.2% in the posttest, as well as a nonsignificant decline in learning from 81.9 % in the pretest, with the effect size being almost negligible at -0.04 SD.

Another major finding was that there was a significant improvement in learning by boys in School 4, as their average score increased from 60% on the pretest to 71.5% in the posttest, with the effect size being 0.90 SD. In Schools 1, 4 and 5, both boys and girls showed improvement in learning, while boys in School 6 and girls in School 3 showed a small decline in learning. Girls in School 6, and both boys and girls in School 2, displayed a significant decline in learning.

To address research objective 2(c) in Subsection 1.5.2, students' achievement was further investigated for all three ability groups in selected classes that displayed contrasting patterns of improvement against a decline. For this, use was made of a frequency distribution of achievement scores, mean and CV. Other things remaining constant, the higher the CV in students' prior knowledge required for learning, the greater is the teaching challenge to deliver the curriculum and, so, greater effort might have to be invested by the class teacher to improve students' learning. Thus, CV reflects the extent of teaching challenge and effort to be invested by teachers to address students' cognitive needs within a class.

For higher CV values of prescores ranging between 0.30 and 0.43, teachers faced relatively increasing challenge in teaching Classes 1, 10 and 12 whereas, for

Class 16, a low value of CV of prescores (0.05) implied that students had the same prior knowledge as assessed in the pretest. For Classes 3, 4, 5, 9, 14 and 15, the extent of teaching challenge was moderate, with their CV values ranging between 0.10-0.20. Students' achievement also varied in these classes.

In Class 7, there were six students in the medium-ability group and 11 students in the high-ability group. The mean postscores of these 17 students declined because eight students scored less than 40% and the remaining nine students scored between 40-60% only. Also, the gap in performance widened between them, as the respective CV value increased to 0.34 in the posttest from 0.24 in the pretest.

In contrast, Classes 1 and 10 had relatively larger numbers of students (21 students in Class 1 and 37 in Class 10). Yet, in Class 1, the postscores increased on average to 57.6% from a low of 46.1% in the pretest, despite a high CV value of 0.43 in the prescore. These teachers of Classes 1 and 10 seemed to have internalised disadvantages of large class size because they divided their students into smaller groups based on prior knowledge, which probably helped them to focus on teaching students according to their levels of ability and paid off because students showed remarkable improvement in their achievement.

In Class 1, there were nine students in the low-ability group. Possibly, these students benefitted from the learning opportunities offered by their mathematics teacher because they achieved higher postscores. Thus, five out of nine students from the low-ability group and five out of seven students from the medium-ability group students moved into the higher range of scoring (60-100%) in the posttest. Consequently, the mean score of this class improved significantly to 57.6%. However, six out of 21 students still needed adequate teacher help.

In Class 10, the students who achieved higher scores of 60-100% in the posttest comprised 6 out of 7 students in the low-ability group and 15 out of 18 in the medium-ability group. Also, all 12 students of the high-ability group continued to score between 60-100%. Thus, all 37 students of Class 10 contributed to a steep increase in mean postscore to 83.9% from an average prescore of 51.1%. As a result, individual differences were also reduced as reflected by a decrease in the CV value of 0.30 in prescores to 0.17 in postscores.

4.6.5 Limitations

Class size as measured in this study excluded many students who were absent for the pretest and / or posttest, but who otherwise attended classes and probably influenced teachers' time, class time and other students' time directly or indirectly. Because they were part of classroom learning environment, their presence in class also impacted other students' learning outcomes, which could not be separated in this analysis and thus remains to be a limitation.

The bivariate correlations estimated in this analysis could be biased upwards because of exclusion of covariables such as teaching experience or class size, which might have confounded the influence of those variables on achievement; this needs to be isolated by partial correlation analysis, as considered in the next chapter.

Chapter 5

MATHEMATICS CLASSROOM LEARNING ENVIROMENT: ITS MEASUREMENT AND ASSOCIATIONS WITH ACHIEVEMENT AND PRIOR KNOWLEDGE

5.1 Introduction

This chapter reports the validation of the new Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) developed and employed in this study to accomplish research objectives 3(a) to 3(d) of Subsection 1.5.2, which include measurement of the construct of classroom learning environment in the mathematics domain, as well as its associations with students' achievement and prior knowledge. This chapter is divided into five sections. Section 5.2 addresses research objective 3(a) involving the development of the MCOLES, including establishing content and face validities, as well as convergent, discriminant and concurrent validities. Section 5.3 addresses research objective 3(b) involving extracting factors of the MCOLES based on exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Section 5.4 deals with research objective 3 (c) by establishing the predictive validity of the MCOLES based on correlations between students' achievement and different MCOLES factors and then, for research objective 3 (d), reports gender differences in correlations for the whole sample and each school. Finally, Section 5.5 offers a chapter summary.

5.2 Development and Validation of MCOLES

This section presents results of development and validation of the MCOLES by following the framework suggested by Trochim and Donnelly (2006), as explained below in four subsections.

5.2.1 Theoretical Rationale for Content Validity and Face Validity of MCOLES

The development of the MCOLES involved establishing translational validity, which comprises content and face validities (Trochim & Donnelly, 2006). For content validity, the new MCOLES was developed after modifying the pre-validated Constructivist Oriented Learning Environment Survey (COLES) (Aldridge et al., 2012). The theoretical basis for the COLES was explained in a detailed literature review in Subsection 2.4.2. Its 11 dimensions and 88 items were closely examined for their theoretical underpinnings and relevance to the mathematics domain. A comparison between items of the MCOLES and COLES was shown in Table 3.3 of Subsection 3.4.3. My version of the MCOLES has 56 items over seven dimensions, therefore requiring a shorter response time from students (see Appendix 8).

The theoretical rationale for the MCOLES, as explained in detail in Subsection 3.4.3, is supported by: theories of mathematics learning including constructivist theory (von Glasersfeld, 1985, 2000); social constructivism (Ernest, 1991, 1998); Moos' (1974, 1979) psychosocial dimensions; the six principles of teaching and learning adopted by the Department of Education & Training in the State Government of Victoria (2012); and the national curriculum authority (NCERT, 2012) in India. This approach provided support for the content validity of the MCOLES.

While developing the MCOLES, a pilot survey administered to a subsample of 20 low-, 20 medium- and 15 high-ability students and their Year 10 mathematics teachers from different schools in this study provided feedback from students and teachers, which established the face validity of items of the MCOLES (Munby, 1997).

5.2.2. Outline of Criterion-related Concurrent and Predictive Validities of MCOLES

Following Trochim and Donnelly (2006), the MCOLES was validated by establishing:

a) criterion-related validity, which includes:

i) convergent validity (or scale reliability) or whether items within a scale are highly correlated with each other

ii) discriminant validity involving whether items from different scales are not highly correlated

b) concurrent validity or whether the instrument can distinguish between groups, or the membership of participants, as expected on theoretical grounds

c) predictive validity or whether the construct can predict what it should theoretically predict, which was assessed by correlations between factors and students' achievement.

Data on students' perceptions were collected by administering the MCOLES to a sample of 531 Year 10 mathematics students and their 11 teachers from six schools, one from Australia and five from India. The rationale for sample selection was explained in Section 1.2 and sample demographics were described in Subsection 3.5.3. After a careful checking, some data gaps were noticed, which accounted for approximately less than 5% of the sample. Missing data were imputed by applying the method of *missing completely at random* (MCAR) (Little & Rubin, 2002; Enders, 2010) as explained in Subsection 3.6.1 and using the M*plus* software (Muthen & Muthen, 2008), which resulted in 511 complete responses to the 56-item MCOLES (see Appendix 8).

5.2.3 Convergent Validity

For establishing convergent validity, a scale reliability analysis was carried out using Cronbach's alpha reliability coefficient by running SPSS (Version 24) on MCOLES responses. Because inter-item correlations could be sensitive at the student level to student-specific differences, comparative estimates were obtained for the student and the class mean as units of analysis. For internal scale consistency, a minimum value of 0.6 is generally accepted as satisfactory (Cronbach, 1951; Yang & Green, 2011). In addition, the mean value of inter-item correlations for each scale was used to identify which combination of items yielded a relatively stronger association (Zwick & Velicer, 1986). Initially, estimates of the alpha coefficient were not satisfactory for many scales except Equity. To improve the scale consistency among items, the following adjustments were made.

Based on inter-item correlations (see Table 5.1), six items (D4, D15, D19, D42, D43 and D49) with low correlations ($r < 0.4$) were dropped. This resulted in higher alpha coefficients for all scales with a range of values between 0.61-0.90 with the student as the unit of analysis, and between 0.78-0.97 when the class mean was used as the unit of analysis.

Also, for Task Orientation by Co-operation and Differentiation, initially estimates of the alpha reliability coefficient were less than the threshold value of 0.6, because D25 and D26 in the former dimension and D41, D44, and D45 in the latter dimension showed a negative correlation with other items. Justified by a meaningful interpretation, they were regrouped. For instance, I created the subgroup (D25, D26, D28) by examining the MCOLES items of Cooperation that students perceived while completing assigned tasks, and the subgroup (D27, D29 to D32) that reflects Task Orientation (see Appendix 8 for the item description).

Results reported in Table 5.1 support a satisfactory range of values of the alpha coefficient for these two subgroups between 0.7 and 0.9 when the class mean was taken as a unit of analysis. So, this regrouping justified splitting this scale into Task Orientation and Cooperation for factor analysis in the next section.

Similarly, after dropping item D42 and D43 which were weakly correlated with other items ($r < 0.4$), two subgroups were formed for Differentiation, justified by a meaningful interpretation. The subgroup (D41, D44, D45) reflects students' perceptions of Task Differentiation, and the subgroup (D46, D47, D48) reflects students' perceptions of Peer Differentiation, while students perceived mathematics tasks assigned to them as different. The estimates of Cronbach's alpha coefficient for these two subgroups were higher in the range 0.7-0.9 with the class mean as the unit of analysis.

This finding also mirrors closely the strategy of grouping students by their levels of ability (prescores) as adopted by teachers in my study. Because of this, factor analysis was carried out by splitting Differentiation into Task Differentiation and Peer Differentiation.

Also, Cronbach's alpha coefficients, estimated separately for 313 boys and 198 girls (Tables A10.1 and A10.2 in Appendix 10), were above the threshold level of 0.6 (Kaiser, 1970, 1974) with either the individual student or the class mean as the unit of analysis. This supports the construct validity of the MCOLES.

5.2.4 Discriminant and Concurrent Validities

The discriminant validity requires that the items within a scale are not highly correlated with items from different scales. This feature was assessed by an analysis of partial correlations between scale means of items with the individual student as the unit of analysis. Partial correlations were used instead of bivariate correlations to minimise the confounding influence of covariables. As there are seven scales in the MCOLES, six partial correlations can be calculated for each scale. If any correlation was nonsignificant, it was treated as being zero, while calculating their mean values over other six scales. The mean partial correlations thus obtained were low (0.08-0.16) for all scales, as they should be, each being less than 0.3 as reported in Table 5.1. These results established discriminant validity of the MCOLES for the overall sample of 511 students.

Similarly, discriminant validity was assessed for boys and girls, separately. The mean values of partial correlations of scales were very low, ranging between 0.09 and 0.17 as shown in Tables 5.2 and 5.3.

Concurrent validity requires that each scale in the MCOLES should distinguish between groups or the membership of students between classes in the sample. For assessing this, the eta^2 statistic which is the ratio of 'between sum of squares' to the 'total sum of squares', was examined by using the class membership as an independent variable, while running the SPSS option of one-way analysis of variance (ANOVA). Table 5.1 shows that estimates of eta^2 were in the range of 0.09-0.19 and were statistically significant at the 1% level for the overall sample Similarly, the concurrent validity of the MCOLES was established for boys and girls, separately, which yielded statistically significant estimates of eta ^{2} between 0.08 – 0.17 for boys and 0.28 – 0.35 for girls (Tables A10.1 and A10.2 in Appendix 10).

Table 5.1: Whole Sample: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=511)

Scale	Unit of Analysis	No. of Items	Cronbach's Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Student Cohesiveness & Personal Relevance after dropping item D4[@]	Student	7	0.700	0.120	0.170*
	Class mean	7	0.913		
Teacher Support after dropping item D15	Student	7	0.847	0.130	0.180*
	Class mean	7	0.941		
Involvement after dropping item D19	Student	7	0.804	0.150	0.160*
	Class mean	7	0.935		
Task Orientation by Co-operation				0.110	0.090*
Co-operation (D25, D26, D28)	Student	3	0.684		
	Class mean	3	0.781		
Task Orientation (D27, D29 to D32)	Student	5	0.655		
	Class mean	5	0.841		

@ The correlation of D4, D15 and D19 with other items were found to be weak ($r < 0.4$).

* $p < 0.01$. The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students.

(contd..)

Table 5.1 (contd..2): Whole Sample: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=511)

Scale	Unit of Analysis	No. of Items	Cronbach's Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Equity	Student	8	0.861	0.160	0.190*
	Class mean	8	0.968		
Differentiation after dropping D42, D43@				0.080	0.090*
Task Differentiation (D41, D44, D45)	Student	3	0.612		
	Class mean	3	0.799		
Peer Differentiation (D46 to D48)	Student	3	0.737		
	Class mean	3	0.895		
Clarity of Assessment Criteria & Feedback after dropping D49@	Student	7	0.805	0.140	0.180*
	Class mean	7	0.968		

@ D42, D43 and D49 were weakly correlated with other items ($r < 0.4$).

* $p < 0.01$. The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students.

5.3 Extraction of Factor Values of MCOLES

This section presents results of extracting factor values of the MCOLES to accomplish my research objective 3(b) of Subsection 1.5.2. The results are divided into four subsections, covering for exploratory factor analysis (EFA) in Subsection 5.3.1 and for confirmatory factor analysis (CFA) of the first-order in Subsection 5.3.2; then, a rationale for a parsimonious solution is described in Subsection 5.3.3 and the uni-factor solution derived by second-order CFA is presented in Subsection 5.3.4.

5.3.1 Factor Extraction of MCOLES by EFA

The factorability of data was assessed by EFA for each scale of the MCOLES using the SPSS software (Version 24) (Pallant, 2013). To investigate the acceptability of estimates, the following three criteria were adopted, as explained in Subsection 3.6.4:

i) Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy should be 0.6 or more.
ii) Bartlett's test of sphericity should be significant.
iii) The correlation between an individual item and other items within the scale must be above 0.3 (Bartlett, 1954; Kaiser, 1970, 1974; Tabachnick & Fidell, 2013).

Table 5.2 provides estimates obtained by EFA for each of the above statistical criteria for the whole sample of 511 students. Evidently, factor analysis was justified for the whole sample of MCOLES responses because the KMO index of sampling adequacy ranged above the threshold value of 0.6 for each scale, from 0.667 for Differentiation to 0.885 for Equity. Also, the Bartlett test of sphericity was satisfactory as indicated by high values of chi-square, ranging from 546.4 for Student Cohesiveness to 1499.3 for Equity, which were all statistically significant at the 5% level. To determine how many factors to extract, I used the criteria of i) the percentage of total variance explained ii) initial eigenvalues iii) scree plots and iv) Monte Carlo Parallel Analysis (PA).

The Monte Carlo PA program (Watkins, 2000) provides estimates of eigenvalues based on randomly-generated data, given the number of items and the number of observations for each scale. These PA eigenvalues were then compared with initial eigenvalues obtained by EFA. The latter eigenvalue was accepted only if it exceeded the corresponding eigenvalue(s) derived from the Monte Carlo PA (Cattell, 1966; Horn, 1965; Hubbard & Allen, 1987).

For the present sample of 511 students, the eigenvalues for all factors were greater than the corresponding values by the Monte Carlo PA for all seven dimensions of MCOLES, which justified the number of factors extracted. The items of Student Cohesiveness and Personal Relevance, Task Orientation by Cooperation, and Differentiation loaded on two factors each, while the items of other four scales loaded on only one factor each. The respective scree plots were shown in Appendix 9. Thus, the latent variable of classroom learning environment perceived by 511 students in terms of a reduced number of 50 items was measured by a 10-factor solution.

Similar to the whole sample, the boys' sample also yielded the same number of factors for each of seven scales (Table A10.3 in Appendix 10). For Differentiation, this yielded two factors with the same grouping of items as considered for the whole sample, while the items of Equity loaded on one factor only.

For the girls' sample, a nine-factor solution emerged because Student Cohesiveness and Personal Relevance yielded only one factor, in contrast to the two factors extracted for boys as well as for the whole sample (Table A10.4 in Appendix 10). For Clarity of Assessment Criteria & Feedback, the girls' sample initially gave two eigenvalue factors, but the second eigenvalue of 1.145 was slightly less than the respective eigenvalue factor from the Monte Carlo PA (1.158). Therefore, only the first factor with the initial eigenvalue of 3.241 was accepted for further analysis. For the girls' sample, the KMO index also ranged above the threshold value of 0.6 for each of seven scales and the item retention criteria suggested the exclusion of the same 6 items (D4, D15, D19, D42, D43 and D49) with low correlations ($r < 0.4$). Table 5.3 reports a summary of results of factor extraction from CFA of MCOLES data, the grouping of items for each factor and estimates of factor loading.

Table 5.2: All Students: Eigenvalues and Factor Determination for MCOLES (N=511)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Student Cohesiveness & Personal Relevance (d=7, by dropping D4)	2.516 (35.9%) 1.248 (17.8%)	1.208 1.125	2	0.756	546.4	21
Teacher Support (d=7, by dropping D15)	3.775 (53.9%)	1.184	1	0.883	1312.1	21
Involvement (d=7 by dropping D19)	3.276 (46.8%)	1.332	1	0.843	984.5	21
Task Orientation by Co-operation (d=8)	2.565 (32.1%) 1.607 (20.1%)	1.182 1.073	2	0.705	768.5	28
Equity (d=8)	4.077 (51.0%)		1	0.885	1499.3	28

(contd..)

Table 5.2 (contd..2): All Students: Eigenvalues and Factor Determination for MCOLES (N=511)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Differentiation (d=6, dropping D42, D43)	2.254 (37.6%) 1.447 (24.1%)	1.182 1.073	2	0.667	586.2	15
Clarity of Assessment Criteria & Feedback (d=7, by dropping D49)	3.250 (46.4%)	1.182	1	0.853	937.3	21

Notes: For obtaining the eigenvalue factors, Principal Axis Factoring and Direct Oblimin rotation were used in all cases except when more than one factor was extracted. In that two-factor case, Varimax rotation was used. Figures within brackets are the respective % of variance explained by items used in extracting a factor.

* Bartlett test of sphericity was found to be significant using chi sq. at the 5% level for all scales.

5.3.2 Factor Determination of MCOLES by CFA

To account for the underlying covariance structure of the MCOLES, I conducted confirmatory factor analysis for all seven dimensions simultaneously after excluding items D4, D15, D19, D42, D43 and D49 (see Appendix 8 for item descriptions) from the original data set of 56 items by using the combination of items that were pre-identified when applying EFA. Using M*plus* software (Version 7.4), CFA was run for the whole sample of 511 students, but not separately for boys and girls, because the combined set of boys and girls represents the covariance structure of underlying co-educational classroom settings in the sample. The CFA yielded maximum likelihood estimates of factor loadings and a 10-factor solution for all 511 students as reported in Table 5.4. Factor loadings for all 50 items ranged between 0.4-0.9 and were statistically significant.

The factor structure of the MCOLES is exhibited in Figure 5.1 with continuous observed variables (d1 – d56) inside square boxes, the latent factors (f 1 – f 7) inside elliptical shapes, and covariant factors joined by arrows. For example, f11 and f12 represent two factors extracted for the first dimension f1, Student Cohesiveness and Personal Relevance. The estimated values of factor loadings shown in the figure correspond to results in Table 5.4. The factors loadings and the covariances between factors were all significant, which confirm the 10-factor solution of the underlying structure of the construct of classroom learning environment (CLE).

Furthermore, the correlations between first-order factors of CFA warrant further investigation for a parsimonious factor representation of CLE as explained in in the next section.

5.3.3 Rationale for a Parsimonious Solution of MCOLES

In pursuit of research objective 3(b) of Subsection 1.5.2, a theoretical rationale is presented for a second-order CFA of MCOLES data in this subsection by including the cross-correlations between first-order factors already extracted (see Table 5.5).

According to Brown (2014), if first-order factors are not theoretically related,

there would be no justification in pursuing higher-order factor analysis. A detailed literature review in Subsection 2.4.2 suggests a theoretical rationale for COLES dimensions' coverage of the three general categories of human environments identified by Moos (1974, 1979). For example, the principle that the learning environment is supportive and productive is mapped to the MCOLES dimensions of Student Cohesiveness and Teacher Support, while the principle that learning connects strongly with communities and practice beyond the classroom is mapped to Moos' notion of Personal Relevance. Also, prior knowledge and its assessment are embedded in the COLES dimensions of Differentiation and Clarity of Assessment Criteria, which can be applied to the mathematics domain. A comparison between the COLES and the MCOLES (Table 3.3), as discussed in Subsection 3.4.3, suggests that six principles of learning and teaching (State Government of Victoria, 2012) underpin a rationale for interrelationships between first-order factors.

Empirically, most factors had strong cross-correlations (Table 5.5), except Task Orientation (F41) and Peer Differentiation (F62), which showed weak correlations with other factors between (0.04-0.25). It can therefore be inferred that students' perceptions of items in those two dimensions were relatively less important compared with other MCOLES factors. Table 5.5 further signals that all other eight factors were closely intercorrelated with estimated values above 0.5. Thus, it is justified empirically to apply second-order CFA to the model variants, as explained in the following subsection.

5.3.4 Second-Order CFA of MCOLES

Basically, two model variants were considered at the second stage:

a) Model 1 included all ten first-order factors of the MCOLES.
b) Model 2 had eight factors only, excluding Task Orientation (F41) and Peer Differentiation (F62), which were weakly correlated with all other factors.

Table 5.3: Factors Extracted and Estimates of Factor Loadings for MCOLES from CFA

Scale	Factor Description	Factor	Items**	Factor Loadings (Range)*
1	a) Student Cohesiveness	F11	D1, D2, D3	0.52-0.63
	b) Personal Relevance	F12	D5, D6, D7, D8	0.55-0.67
2	Teacher Support	F2	D9 TO D14, D16	0.60-0.74
3	Involvement	F3	D17, D18, D20 TO D24	0.51-0.73
4	a) Cooperation	F41	D25 to D28	0.40-0.89
	b) Task Orientation	F42	D27, D29 to D32	0.46-0.56
5	Equity	F5	D33 to D40	0.58-0.73
6	a) Task Differentiation	F61	D41, D44, D45	0.48-0.73
	b) Peer Differentiation	F62	D46 to D48	0.57-0.83
7	Clarity of Assessment Criteria & Feedback	F7	D50 to D56	0.45-0.69

Note: * Maximum likelihood estimates from CFA were all statistically significant at 1% level.
** For the item description, see Appendix 8.

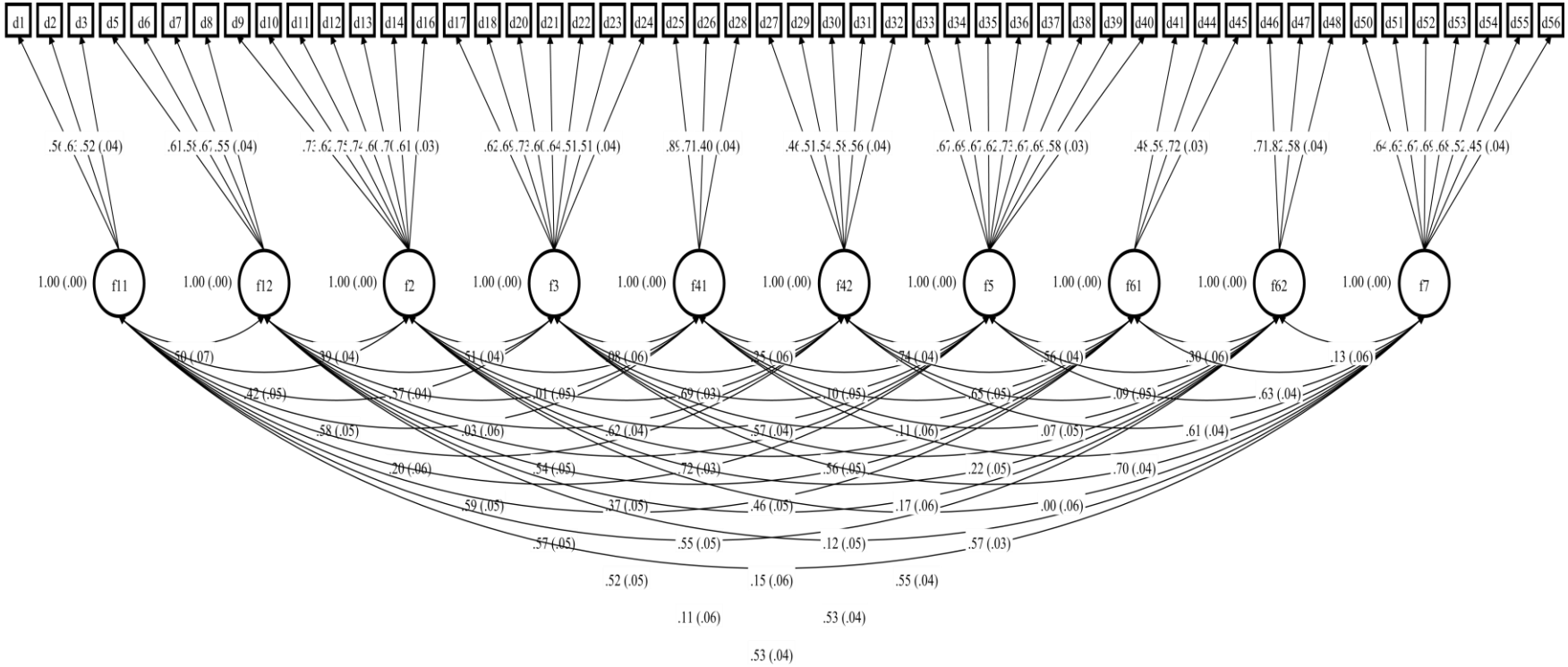


Figure 5.1: A Ten-Factor Solution of MCOLES Structure with 50 Items, N=511

Notes: See Appendix 8 for item description. Factor loadings shown above along the arrows are also given in Table 5.4, while covariances between the latent factors are shown by connecting arrows. All estimates are statistically significant at the 5 % level.

Table 5.4: Maximum Likelihood Estimates of Factor Loadings from CFA for MCOLES Responses (N=511)

Classroom Learning Environment Factors	Factor loadings for MCOLES									
	F11	F12	F2	F3	F41	F42	F5	F61	F62	F7
Student Cohesiveness (F11)										
D1	0.557									
D2	0.634									
D3	0.524									
Personal Relevance (F12)										
D5		0.609								
D6		0.577								
D7		0.668								
D8		0.553								
Teacher Support (F2)										
D9			0.731							
D10			0.618							
D11			0.749							
D12			0.738							
D13			0.601							
D14			0.703							
D16			0.614							

(contd..)

Table 5.4 (contd..2): Maximum Likelihood Estimates of Factor Loadings from CFA for MCOLES Responses (N=511)

Classroom Learning Environment Factors	Factor loadings for MCOLES									
	F11	F12	F2	F3	F41	F42	F5	F61	F62	F7
Involvement (F3)										
D17				0.618						
D18				0.694						
D20				0.729						
D21				0.602						
D22				0.644						
D23				0.511						
D24				0.511						
Task Orientation (F41)										
D25					0.887					
D26					0.713					
D28					0.402					
Co-operation (F42)										
D27						0.464				
D29						0.513				
D30						0.535				
D31						0.576				
D32						0.556				

(contd...)

Table 5.4 (contd..3): Maximum Likelihood Estimates of Factor Loadings from CFA for MCOLES Responses (N=511)

Classroom Learning Environment Factors	Factor loadings for MCOLES									
	F11	F12	F2	F3	F41	F42	F5	F61	F62	F7
Equity (F5)										
D33							0.667			
D34							0.687			
D35							0.667			
D36							0.616			
D37							0.727			
D38							0.669			
D39							0.686			
D40							0.579			
Task Differentiation (F61)										
D41								0.588		
D44								0.722		
D45										
Peer Differentiation (F62)										
D46									0.707	
D47									0.821	
D48									0.577	

(contd...)

Table 5.4 (contd..4): Maximum Likelihood Estimates of Factor Loadings from CFA for MCOLES Responses (N=511)

Classroom Learning Environment Factors	Factor loadings for MCOLES									
	F11	F12	F2	F3	F41	F42	F5	F61	F62	F7
Clarity of Assessment Criteria & Feedback (F7)										
D50										0.636
D51										0.629
D52										0.671
D53										0.689
D54										0.683
D55										0.517
D56										0.450

Notes: Factor loadings were estimated by CFA using the M*plus* software (Version 7.4). All estimates were significant at the 1 % level.
Six items: D 4, D15, D19, D42, D43 and D49 were dropped from analysis because each item correlation was less than 0.4.
For description of items, see Appendix 8.

Table 5.5: Cross-Correlations between First-Order Factors Extracted by CFA for MCOLES

Classroom Learning Environment Factors	Cross-Correlations between First-Order Factors Extracted from CFA of MCOLES									
	F11	F12	F2	F3	F41	F42	F5	F61	F62	F7
Student Cohesiveness (F11)	1.000									
Personal Relevance (F12)	0.608	1.000								
Teacher Support (F2)	0.598	0.519	1.000							
Involvement (F3)	0.690	0.630	0.602	1.000						
Task Orientation (F41)	0.156	0.065	0.041	0.095	1.000					
Co-operation (F42)	0.780	0.684	0.741	0.791	0.162	1.000				
Equity (F5)	0.704	0.544	0.737	0.671	0.108	0.834	1.000			
Task Differentiation (F61)	0.692	0.636	0.621	0.691	0.119	0.813	0.711	1.000		
Peer Differentiation (F62)	0.156	0.160	0.140	0.181	0.248	0.165	0.135	0.244	1.000	
Clarity of Assessment Criteria & Feedback (F7)	0.684	0.618	0.641	0.673	0.047	0.817	0.710	0.732	0.157	1.000

Maximum likelihood (ML) estimates of factor loadings and their respective R^2 values are shown for both models in Table 5.6. The ML estimates of all factor loadings varied between 0.6 to 0.9 for both options, except for Task Orientation (0.116) and Peer Differentiation (0.175), which had low R^2 values that were statistically nonsignificant.

Table 5.6: Factor Loadings Obtained by Second-Order CFA

Factors Loading on the Construct, Classroom Learning Environment	Model 1		Model 2	
	ML Ests.*	R^2	ML Ests.	R^2
Student Cohesiveness (F 11)	0.693	0.480	0.692	0.479
Personal Relevance (F12)	0.619	0.383	0.619	0.383
Teacher Support (F2)	0.721	0.520	0.722	0.521
Involvement (F3)	0.753	0.567	0.753	0.563
Co-operation (F42)	0.889	0.789	0.888	0.789
Equity (F5)	0.809	0.654	0.809	0.654
Task differentiation (F61)	0.742	0.550	0.741	0.549
Clarity of Assessment Criteria & Feedback (F7)	0.701	0.611	0.782	0.611
Task Orientation (F41)	0.116	0.014		
Peer Differentiation (F 62)	0.175	0.031		

Note: * All estimates above were obtained by using M*plus* (Version 7.4) and were statistically significant at the 1% level, except for factors F41 and F62, which were therefore ignored in Model 2.

Interestingly, even the exclusion of F41 and F62 led to little change in the proportion of variance explained by the remaining factor indicators in Model 2. For Model 2, R^2 values were relatively high, ranging from 38.3% for Personal Relevance (F12) to as high as 78.9% for Co-operation (F42), and were statistically significant at the 1% level.

For a comparative evaluation of Model 1 and Model 2, the discussion about model fit indices in Subsection 3.7.4 suggests that, among the absolute fit indices, the SRMR is generally preferred over the chi-square value (Brown, 2014). For a model fit to be acceptable, the value of SRMR should be as close to zero as possible. Simulation studies by Hu and Bentler (1999) suggest a value of 0.08 or below for reasonably good fit, with 0.0 indicating perfect fit.

Among the parsimony correction indices, the RMSEA index is generally preferred over others, and its value should be as small as possible and close to zero, which indicates a perfect fit.

Among comparative fit indices, CFI or TLI can be considered. In either case, the value of the index should be as close to 1 as possible. Values in the range of 0.9 and 0.95 are indicative of an acceptable model fit (Bentler, 1990).

In the present case, among the absolute fit indices shown in Table 5.7, the chi-square value was high at 2001 with 1165 degrees of freedom for Model 1, and at 2071 with 1171 degrees of freedom for Model 2. Both cases support the alternative

Table 5.7: Results of Second-Order CFA of MCOLES and Model Fit Information

Model Fit Criteria	Model 1	Model 2
Log L	-36026	-36019
Chi-square	2001	2071
Degrees of Freedom	1165	1164
SRMR	0.062	0.062
RMSEA	0.039	0.039
CFI	0.872	0.873
TLI	0.865	0.866
AIC	72372	72361

hypothesis that the model estimates do not reproduce the sample variances and covariances. That is, the model does not fit the data well. Use of chi-square values to determine model fit is subject to criticism as it often far exceeds the critical value at given degrees of freedom, which was also true in the present case (Brown, 2014; Hu & Bentler, 1995, 1998, 1999).

Alternatively, the SRMR value was around 0.06 for both models, which supports good model fit according to Hu and Bentler (1999).

The parsimony correction index, RMSEA, was also very low at 0.039 for both models, which indicates a reasonably good model fit. The comparative fit index, CFI (or TLI), was slightly higher and better for Model 2 at 0.873 than for Model 1 at 0.872.

Moreover, the AIC value was less for Model 2 (72361) than for Model 1 (72372). Hence, Model 2 is preferable.

Finally, the uni-factor solution extracted from Model 2 based on the evaluation criteria offers a reasonably good fit for the construct of classroom learning environment, as exhibited in Figure 5.2. Thus, the factor values extracted from Model 2 were considered for investigating predictive ability of the MCOLES by correlation analysis in the next section.

5.4 Predictive Validity of MCOLES by Correlation Analysis

This section presents results for the predictive validity of the MCOLES and addresses research objectives 3(c) and 3(d) of Subsections 1.5.2. First, an analysis of correlations between students' achievement (postscores) and each MCOLES factor is reported in Subsection 5.4.1using the factor values derived by CFA for the whole sample and each class, separately. Then, results of gender differences in correlations are presented in Subsection 5.4.2 for the whole sample and each school, separately.

5.4.1 Predictive Validity of MCOLES by Correlations for Whole Sample, Each School and Each Class

In this subsection, the predictive validity of MCOLES is assessed by examining the correlations between each MCOLES factor and students' achievement for the overall sample, each school and each class.

In a preamble to administering the MCOLES, students were requested to respond to items of the MCOLES soon after their teachers taught them a topic of measurement from the Year 10 mathematics curriculum. The MCOLES responses thus reflect students' perceptions of mathematics classroom learning environment just experienced by them.

Out of 511 students who responded to the MCOLES, 47 students did not participate in the pretest or / and posttest. The factor values for 464 students (285 boys

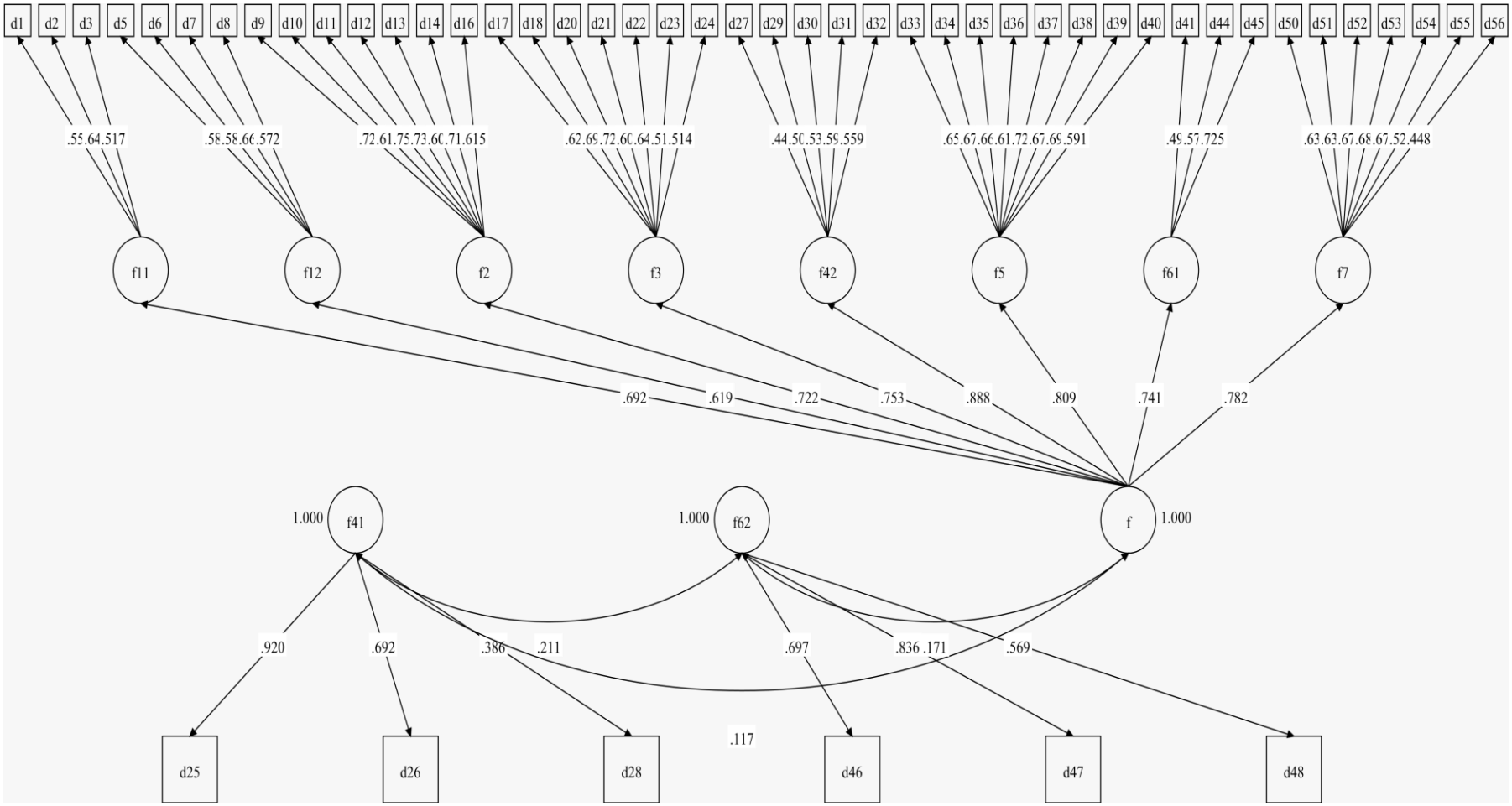


Figure 5.2: Second-Order CFA of MCOLES

Note: At the first stage, ten factors f11 to f7 were extracted by CFA using 50-item MCOLES with N =511 as shown in Figure 5.1

and 179 girls) who complied with all survey requirements were therefore used for computing correlations between students' achievement and classroom learning environment factors (Table 5.8).

For the whole sample (see the second-last column of the last row of Table 5.8), the correlation between students' achievement and the overall classroom learning environment (taking all factors combined) was 0.23 and statistically significant.

The bivariate correlation (the last row of Table 5.8) between achievement and each MCOLES factor also was significant for as many as eight out of ten factors and varied from 0.10 for Personal Relevance to 0.28 for Teacher Support, but was nonsignificant for Task Orientation and Peer Differentiation.

For individual classes in different schools, correlations between achievement and the overall classroom learning environment (all factor combined) (see the second-last column of Table 5.8) varied from 0.37 for Class 3 to 0.60 for Class 9. These correlations were significant only for Classes 3, 4, 8, 9 and 15.

Student Cohesiveness was strongly correlated with achievement at 0.60 for Class 15, whereas it was nonsignificant for many classes. Similarly, Involvement and Cooperation were significantly correlated with achievement in five out of 15 classes, while Personal Relevance was not significantly correlated with achievement for almost all classes except Class 4 (r = 0.38).

In contrast, correlations between achievement and i) Teacher Support in Class 1 and ii) Task Orientation in Class 5 were negative at -0.48 and -0.55, respectively. The negative correlations imply two possibilities because the mean postscores of students indicated an improvement in Class 1, but a decline in Class 5 (see Table 4.10). Many students of Class 1 improved their performance in learning, even though they did not perceive favourable support from their teacher. Also is possible that some students who did not perform well on the posttest perceived a favourable level of teachers' encouragement.

Similarly, in Class 5, many students had low mean postscore of 31.4 % (Table 4.10) even though they perceived a favourable level of Task Orientation. In contrast,

Table 5.8: Correlations of Students' Achievement with MCOLES Factors and Prior Knowledge for Each Class

School	Class	Correlations with Achievement											
		Student Cohesive -ness	Personal Relevance	Teacher Support	Involve ment	Task Orientation	Cooper ation	Equity	Task Differentia -tion	Peer Differentia -tion	Clarity of Assessment Criteria & Feedback	Classroom Learning Environment (All factors)	Prior knowledge
1	1			-0.48*									0.56**
2	3	0.33*			0.37*		0.36*	0.36*	0.44**			0.37**	
	4	0.37*	0.38*		0.33*		0.36*	0.35*	0.40**	0.31*	0.32*	0.42**	0.38*
	5					-0.55*							
3	6							0.31*					
	8	0.45**		0.37*	0.36*		0.45**	0.35*			0.41**	0.43**	
4	9	0.41*		0.48**	0.59**		0.62**	0.53**			0.57**	0.60**	0.75**
	10												0.53**
5	11												0.55**
	12												0.60**
6	14												0.42*
	15	0.60**		0.56**	0.53**		0.47**	0.63**				0.55**	
Sample		0.15*	0.10*	0.28*	0.16*		0.21*	0.23*	0.19*		0.21*	0.23*	0.35*

Notes: Correlations for classes shown above were tested for significance by using online tools (Lowry, 2017).
*$p < 0.05$, ** $p < 0.01$

Table 5.9: Partial Correlations between Students' Achievement and Each MCOLES Factor with Control Variables: Prior Knowledge, Teaching Experience and Class Size

MCOLES Factors	Factor	Partial Correlations with Achievement		
		All	Boys	Girls
Student Cohesiveness	F11	NS	0.15	NS
Personal Relevance	F12	NS	NS	NS
Teacher Support	F2	0.23	0.30	0.15
Involvement	F3	NS	0.12	NS
Task Orientation	F41	NS	-0.15	NS
Co-operation	F42	0.13	0.17	NS
Equity	F5	0.16	0.23	NS
Task Differentiation	F61	0.09	0.17	NS
Peer Differentiation	F62	NS	NS	NS
Clarity of Assessment Criteria & Feedback	F7	NS	NS	NS

Notes: The above correlation coefficients were statistically significant at the 5% level unless indicated by NS.
The estimates were calculated by using 461 observations (283 boys, 178 girls) because, out of 464 who complied with all survey requirements, three observations were outliers, which were excluded.

students' prior knowledge and their achievement were moderately and significantly correlated in as many as seven classes, varying from 0.38 for Class 4 to 0.75 for Class 9 (see the last column of Table 5.8). Also, for the whole sample, the correlation between students' prior knowledge and achievement, which was 0.35 and significant, was stronger than any classroom learning environment factor (last row of Table 5.8).

The bivariate correlations discussed above could be biased upwards because of the confounding influence of covariables, including students' prior knowledge, years of teaching experience and class size, on students' achievement (Kendall et al., 1973). For minimising the confounding influence, partial correlations were estimated using SPSS (Version 24) as reported in Table 5.9. Surprisingly, when the above covariables were controlled, the only significant partial correlations were those between achievement and each of the four classroom factors of Teacher Support (0.23), Equity (0.16), Cooperation (0.13) and Task differentiation (0.08). For the whole sample, of all partial correlations estimated, the one between Teacher Support and achievement was the highest at 0.23, which is comparable with the corresponding bivariate

correlation of 0.28. But the partial correlation between achievement and each of the remaining six factors was nonsignificant.

5.4.2 Gender Differences in Correlations for Whole Sample and Each School

This subsection addresses research objective 3(d) of Subsection 1.5.2 and presents gender differences in correlations between students' achievement and each MCOLES factor, the overall variable of classroom learning environment (CLE) and students' prior knowledge.

First, gender differences in bivariate correlations were tested for statistical significance using the Fisher's *r*-to-*z* transformation (Lowry, 2017) as mentioned in Subsection 3.6.3. Correlations and gender differences in correlations are reported in Table 5.10.

Student Cohesiveness and Personal Relevance only, and nonsignificant for other factors. For the whole sample and all factors of classroom learning environment combined, the gender difference in correlation between students' achievement and classroom learning environment was nonsignificant. However, correlations with achievement were somewhat higher for boys than girls for Teacher Support, Cooperation and Equity. For individual schools, however, there existed gender differences in bivariate correlations in Schools 2, 4 and 6. In School 3, the classroom learning environment for all factors combined was moderately correlated with achievement at 0.47 for boys and at 0.35 for girls, but the gender difference in correlation was not statistically significant at the 5% level. In Schools 2 and 6, however, the correlation between achievement and classroom learning environment was stronger for boys than for girls whereas, in School 4, it was 0.42 for girls but nonsignificant for boys.

In contrast, in Schools 1 and 5, correlations between classroom learning environment and students' achievement were very weak and statistically nonsignificant.

Referring to Table 5.9, gender differences in partial correlation were significant in contrast with those for bivariate correlations, which is attributable to the extent of

Table 5.10: Bivariate Correlations of Students' Achievement with Each MCOLES Factor, and Gender Differences in Correlations between Students' Achievement and All MCOLES Factors Combined in Schools

MCOLES Factors / Schools	Correlations with Achievement			Gender Differences in Correlations
	All (N=464)	Boys (N=285) r_b	Girls (N=179) r_g	$r_b - r_g$
MCOLES Factors				
Student Cohesiveness	0.15	0.21	NS	0.21
Personal Relevance	0.10	NS	0.15	-0.15
Teacher Support	0.28	0.32	0.26	NS
Involvement	0.16	0.16	0.22	NS
Task Orientation	NS	NS	NS	NS
Cooperation	0.21	0.25	0.22	NS
Equity	0.23	0.27	0.23	NS
Task Differentiation	0.19	0.22	0.18	NS
Peer Differentiation	NS	NS	NS	NS
Clarity of Assessment Criteria & Feedback	0.21	0.25	0.21	NS

Schools	Correlations between Achievement and Classroom Learning Environment (All factors combined)			Gender Differences in Correlations
	ALL (N= 464)	Boys (N=285)	Girls (N=179	
All Schools	0.23	0.28	0.21	NS
School 1	NS	NS	NS	NS
School 2	0.24	0.32	NS	0.32
School 3	0.44	0.47	0.35	NS
School 4	0.22	NS	0.42	-0.42
School 5	NS	NS	NS	NS
School 6	0.41	0.32	NS	0.32

Note: * Correlations shown above were all statistically significant unless indicated by NS.
Gender differences in correlations were tested for significance by using the Fisher's *r*-to-*z* transformation (Lowry, 2017).

confounding of covariables on achievement. For example, the partial correlationbetween achievement and Teacher Support was higher for boys at 0.30 than girls (0.15). Comparatively, the bivariate correlations were 0.38 for boys and 0.30 for girls, which implies that gender differences were significant for partial correlations, but nonsignificant for bivariate correlations. As expected, partial correlations were smaller than the respective bivariate correlations and, hence, gender differences were also different in magnitude. Of all factors, Task Differentiation had the lowest partial correlation with achievement at 0.09 for the overall sample, but the gender difference was significant because the partial correlation between Task Differentiation and

achievement was relatively higher and significant for boys at 0.17 but nonsignificant for girls.

5.5 Chapter Summary

The findings for my third research question are summarised in three subsections with i) development and validation of the new MCOLES for measuring the construct of classroom learning environment in Subsection 5.5.1 ii) correlation analysis of predictive validity of the MCOLES in Subsection 5.5.2 and iii) gender differences in correlations between students' achievement and classroom learning environment factors in Subsection 5.5.3.

5.5.1 Validation of MCOLES

Developing and validating the new MCOLES involved a two-stage approach that was guided by Trochim and Donnelly's (2006) framework to ensure that both the translation and criterion-related validity requirements were fulfilled. Translation validity includes content validity which focuses on theoretically-sound representations of the constructs, and face validity emphasises clear interpretations of the items by participants.

At stage 1, an extensive review of related literature provided a rationale including constructivist theory for mathematics learning, the philosophy of social constructivism, Moos' psychosocial environments, six principles of teaching and learning adopted by the Department of Education, State Government of Victoria, and the underlying principles of the national curriculum designed by the NCERT (2012) in India. These supported a more-concise seven-dimensional 56-item MCOLES for mathematics classes after modifying a pre-validated 11-dimensional 88-item COLES. The content validity was thus established. As recommended by Munby (2006), face validity was justified by the feedback from a pilot survey of teachers and students from participating schools regarding improving certain items to enhance meaningful student interpretation.

Validation of the MCOLES involved 511 students and 11 teachers of Year 10 mathematics classes in six schools, including one from Australia and five from India. In a preamble to administering the MCOLES, students were informed of the purpose of the survey. Because responses were collected soon after students were taught a mathematics topic from their respective curricula, items of the MCOLES referred to their most recent experiences and perceptions of mathematics classroom learning environment as provided by their teachers.

MCOLES scale reliability was based on Cronbach's alpha coefficient with both the individual student and the class mean as the unit of analysis. Six items (D4, D15, D19, D42, D43 and D49) (see Appendix 8 for a description) were dropped from further analysis as they were very weakly correlated ($r < 0.4$) as recommended by Kaiser (1970, 1974). This improved reliability for each scale.

In addition, because some items in Task Orientation by Cooperation and Differentiation were negatively correlated with the other items, they were regrouped meaningfully as i) Task Orientation and Cooperation and ii) Peer Differentiation and Task differentiation, which improved the reliability for these two subgroups.

Discriminant validity of the MCOLES was established by examining partial correlations between MCOLES scales. Partial correlations were considered instead of bivariate correlations for minimising the extent of confounding influence of items. The mean partial correlations of scales were very low and in the range of 0.08 – 0.16.

The concurrent validity of MCOLES was established using one-way ANOVA. The eta^2 ratio was statistically significant for all scales and in the range of 0.08-0.35, thereby establishing the validity of the MCOLES not only for the whole sample, but also for the boys' subgroup (N=313) and the girls' subgroup (N=198) separately.

At the second stage, research objective 3(b) of Subsection 1.5.2 was accomplished by applying EFA to the seven dimensions of the MCOLES and considering combinations of items that were pre-identified in the validation process. The factorability of the MCOLES was justified by Bartlett's test of sphericity which was statistically significant at the 5% level, and the Kaiser-Meyer-Olkin (KMO) index of sampling adequacy which was also above the critical level of 0.6 for all scales. The

EFA resulted in a ten-factor solution, extracting one factor each for Teacher Support, Involvement, Equity and Clarity of Assessment Criteria & Feedback; and two factors each for Student Cohesiveness and Personal Relevance*;* Task Orientation by Cooperation*;* and Differentiation. These ten factors also justified a meaningful interpretation of items that were subgrouped in the process of validation and supported by scree plots for the extraction of ten factors for the whole sample of 511 students. and the boys' group of 313 observations. However, only nine factors were extracted for the girls' group of 198 observations because Student Cohesiveness and Personal Relevance loaded on one factor.

Then, to account for the covariance structure of students in co-educational settings, the CFA was applied to MCOLES responses for the whole sample of 511 students to all seven dimensions simultaneously using the M*plus* software. This resulted in a ten-factor solution with all maximum likelihood estimates being statistically significant. These results strongly supported the MCOLES' ten-factor covariance structure as displayed in Figure 5.1.

Also, supported by a theoretical rationale, a careful review of cross-correlations between these ten factors indicated that there was a moderate to strong interrelationships among the first-order factors, except for Task Orientation and Peer Differentiation for which correlations were weak (0.04-0.24), thereby warranting further investigation for a parsimonious factor solution. Then, second-order CFA yielded a single-factor solution for the construct of classroom learning environment (CLE).

A single-factor representation of CLE was found to be amenable for establishing the predictive validity by an analysis of correlations between achievement and MCOLES factors, both individually and jointly as a single factor of CLE.

5.5.2 Predictive Validity by Analysis of Correlations

When predictive validity was tested using Pearson correlations between each of the MCOLES factors and students' mathematics achievement, these correlations

were statistically significant. These results can be used to guide teachers in improving their teaching practice using the MCOLES as a classroom tool.

A main finding for the whole sample was that the correlation of students' achievement with eight out of 10 factors was statistically significant, but weak to moderate in magnitude (between 0.12 and 0.28). The correlations of achievement with Task Orientation and Peer Differentiation were nonsignificant. The predictability of students' achievement from the classroom learning environment factors varied to some extent among the ten factors extracted, reflecting their relative importance.

Another major finding of this study was that Teacher Support had the strongest association with achievement for both boys' and girls' subgroups, separately. The correlations were statistically significant at 0.33 for boys and 0.25 for girls, which could be biased upwards due to a possible confounding influence of covariables.

The partial correlations which minimise such confounding influences were estimated after controlling for covariables including students' prior knowledge, years of teaching experience and class size. Partial correlations with achievement were slightly lower at 0.30 for boys and 0.15 for girls and were statistically significant at the 5% level. For girls, the partial correlation was significant only for Teacher Support and was nonsignificant for all other factors.

For individual classes, too, the correlation of Teacher Support with achievement was moderately higher for two Classes 8 and 9 at 0.37 and 0.48, respectively, but was nonsignificant for other classes.

In Class 5, the correlation between students' achievement and Task Orientation was negative and significant. Because the mean postscore for this class was very low at around 31%, it seems that many students did not perform well despite having favourable Task Orientation.

5.5.3 Gender Differences in Correlations

To accomplish research objective 3(d) of Subsection 1.5.2, bivariate correlations between students' achievement and classroom learning environment were

estimated for each individual factor, as well as for all factors combined for the whole sample and for each of six schools, by using the formula suggested by Snedecor and Cochran (1989) and Lowry (2017).

For the whole sample, taking all MCOLES factors combined, there was no significant gender difference in bivariate correlation between students' achievement and classroom learning environment.

For the boys' group, correlations between achievement and eight of the MCOLES factors were statistically significant and varied between 0.33 and 0.54, with correlations being stronger for Student Cohesiveness followed by Equity, Task Differentiation and Clarity of Assessment Criteria & Feedback. They were nonsignificant for Task Orientation and Peer Differentiation. In contrast, for the girls' group, the correlation between achievement and Personal Relevance was stronger than for other scales including Involvement and Co-operation.

Small gender differences in correlations with achievement were found only for Student Cohesiveness and Personal Relevance, which were statistically significant at the 5% level , but correlations were nonsignificant for other eight factors.

Gender differences were significant in Schools 2, 4 and 6, but nonsignificant in Schools 1 and 3. In Schools 2 and 6, the correlation between classroom learning environment and achievement was stronger for boys than for girls whereas, in School 4, the correlations with achievement was stronger for girls than for boys. In School 3, the bivariate correlations between all classroom learning environment factors combined and achievement was moderate for boys at 0.47 and girls at 0.35, but gender differences in these correlations were not statistically significant at the 5% level.

However, when partial correlations were considered between achievement and all MCOLES factors, gender differences were significant, which can be attributed to the confounding influence of covariables on the achievement of boys and girls. Between Teacher Support and achievement, the partial correlation was higher for boys (0.30) than girls (0.15). For most of MCOLES factors, gender differences were significant because partial correlations were significant for boys and nonsignificant for girls.

The next chapter deals with the role of students' self-efficacy beliefs in mediating the influence of students' prior knowledge and classroom learning environment on achievement.

Chapter 6

STUDENTS' MATHEMATICS SELF-EFFICACY BELIEFS: MEASUREMENT AND THEIR ASSOCIATIONS WITH ACHIEVEMENT, PRIOR KNOWLEDGE AND CLASSROOM LEARNING ENVIRONMENT

6.1 Introduction

This chapter addresses research objectives 4(a) to 4(d) of Subsection 1.5.2 involving the role of self-efficacy in students' achievement. For accomplishing research objective 4(a), Section 6.2 describes the development of the Mathematics Self-Efficacy Scale (MSES) for the measurement of self-efficacy beliefs in achieving mathematics tasks. For examining research objective 4(b), Section 6.3 provides the calibration of students' self-efficacy judgements using MSES responses to estimate their expected scores as a measure of self-efficacy expectancies in achieving tasks and, hence, to identify underachievers as well as students with lower self-efficacy expectancies. Section 6.4 deals with research objective 4(c) involving the predictive validity of the MSES by means of a correlation analysis of efficacy expectancies with students' a) prior knowledge, b) classroom learning environment and c) achievement. To address research objective 4(d), Section 6.4 also reports gender differences in efficacy- expectancies and their correlations with achievement. Section 6.5 presents a chapter summary.

6.2 Development of MSES

The development of the MSES was based on resource material provided to students when seeking their responses to tasks on the posttest according to a guide recommended by Bandura (2006). This section explains the development and validation of 31 mathematics tasks on the posttest.

The content validity of the tasks in the resource material was underpinned by a theoretical rationale that adopted the assessment criteria of equity and fairness. Different tasks were designed according to the knowledge framework, recommended by Hailikari, Nevgi and Lindblom-Ylänne (2007) and discussed in Subsection 3.4.4, involving a mathematics topic of measurement that teachers selected for this study and taught at the Year 10 level after assessing students' prior knowledge of this topic.

As the survey was administered in Australia and India, tasks were drawn from the respective curricula prescribed for the Year 10 level in these countries. There were minor differences in weights assigned to tasks (see Table 3.2) based on the curricula implemented by teachers in their respective classes, as shown in resource material for Class 1 (Appendix 4 for Australia) and Classes 3-16 (Appendix 6 for India).

The 31 tasks on the posttest were classified under the following types.

D1: Declarative knowledge to recall facts

D2: Declarative knowledge to recall concepts and meanings

P1: Procedural knowledge for integration of ideas

P2: Procedural knowledge for solving application tasks.

When the test content was reviewed by Year 10 mathematics teachers, it was found necessary to further classify the tasks in an increasing order of difficulty from the students' perspective, which resulted in creating two more types:

P3: Procedural knowledge for solving mixed tasks

P4: Procedural knowledge needed for higher-order thinking and solving challenging tasks.

Thus, the development of the tasks established the content validity of the MSES.

For the face validity of resource material, feedback was obtained, following Munby (1997), from survey respondents for the evaluation of each item. The teachers in the study ensured that the wording used in the resource material was familiar to students. Moreover, without loss of confidentiality for testing, teachers used similar practice tasks to ensure that their students gained an understanding of solving such tasks. Also, to ensure equity and fairness of testing, students were provided with the assessment criteria and a marking scheme, as shown in Table A9.1 in Appendix 9.

After considering their classroom experiences, teachers provided useful feedback and suggestions to the researcher for rephrasing and modifying some of the tasks proposed. The feedback from students also confirmed their understanding of tasks. The test material therefore was consistent and fair for seeking suitable responses to the MSES items. These procedures helped in establishing the face validity of the MSES and accomplished research objective 4(a) of Subsection 1.5.2

6.3 Assessment of Students' Self-Efficacy Expectancies by Bandura's Method

This section addresses research objective 4(b) of Subsection 1.5.2 by calibrating students' self-efficacy judgements and estimating students' efficacy expectancies in terms of scores expected by them (expected scores) following the method recommended by Bandura (2006*)*. The purpose was to establish if the MSES could be used as a teacher's tool for identifying underachievers or those students with lower self-efficacy expectancies. To do so, the scores expected by students were compared with the postscores awarded by their teachers after evaluation. Subsection 6.3.1 explains the calibration procedure adopted with an example of responses provided by a student in the survey, while Subsection 6.3.2 reports my analysis of efficacy expectancies and pertinent causes of underachievement.

6.3.1 Calibration of Self-Efficacy Judgements into Expected Scores

In a preamble to the survey, students were guided by teachers to follow the procedures for giving responses to various items on the MSES. Teachers clarified students' questions by explaining with examples as given below, so that students' responses would reflect self-efficacy judgements of their own capabilities to solve given tasks on the posttest.

Bandura (2006) recommended a method of calibration of self-efficacy judgements into expected scores (termed students' efficacy-expectancies), which are not the *outcome expectancies* as discussed by previous researchers (Davies & Yates, 1982; Maddux, Norton, & Stolenberg, 1986).

For seeking responses, first, students were asked to read the resource material but not to work out an answer to any test question (task). Teachers guided students concerning how a specific test item with a serial number from the resource material (see Appendices 4 for Class 1 and 6 for Classes 3-16) could be matched with that given in the MSES response format (Appendices 5 for Class 1and 7 for classes 3-16), and then how to give a response to that item by circling an appropriate choice on a scale of 1 to 5, as shown in Figure 3.2 of Subsection 3.4.4. This procedure, following Bandura (2006), clarified what scores students could expect for each response to an item selected from the MSES. Later, all students responded to the posttest under strict test conditions. Teachers scored students' answers according to a marking scheme shown in Table A9.1 in Appendix 9 to determine postscores which were shared with the researcher. For example, considering a student's MSES responses, the procedure for converting responses to equivalent scores is shown in Table 6.1.

Table 6.1: Self-Efficacy Judgements and Expected Scores as Revealed by MSES Responses: An Example of Calibration to Identify Underachieving Student 101011

Item Code	Student's Responses	Expected Score	Item Code	Student's Responses	Expected Score
D11	2	25	P201	4	75
D12	2	25	P202	4	75
D13	1	0	P203	4	75
D14	1	0	P204	3	50
D21	5	100	P205	3	50
D22	5	100	P206	3	50
D23	5	100	P301	3	50
D24	5	100	P302	3	50
D25	4	75	P303	2	25
D26	4	75	P304	3	50
P101	2	25	P305	3	50
P102	2	25	P401	3	50
P103	3	50	P402	2	25
P104	3	50	P403	5	100
P105	3	50	P404	3	50
			P405	4	75

Total score expected by Student 101011 on the posttest	54.8
Total score awarded by the teacher (postscore)	22.0
Underachievement estimated as % of postscore	149.1

Each item on the MSES was assigned a code to identify the types of tasks considered according to the knowledge framework (Hailikari, Nevgi, & Lindblom-

Ylänne, 2007)). For example, Task Type D1 has four questions on the MSES, with item codes of D11 to D14, for which the student's responses were converted into expected scores according to the procedure shown below.

For items D11 and D12, the student's response was 2, which means that the student was unsure how to solve that task, which was therefore converted to an expected score of 25%. Similarly, for items D13 and D14, the student's response was 1, which was converted to an expected score of 0%. The total score expected for all four items of task type D1 was therefore an average of 12.5% only (=50/4). Similarly, the total score expected for all 31 items together was on average 54.8%.

But the postscore that the teacher awarded to the student was only 22.0%, after the evaluation of student's mistakes, possibly because this student over-assumed self-judgemental capabilities in solving the tasks on the posttest. Thus, this student was an underachiever, with the deviations being estimated at 149.1% (= 100 x (54.8 – 22.0) / 22.0).

By applying a similar procedure to the whole sample, the scores expected by all students were derived and then compared with their postscores. The deviations as a percentage of postscores were calculated. A deviation of ± 5% was allowed as a margin of error in matching the expected scores with the postscores.

When the deviation was more than 5%, students were identified as underachievers (see Appendix 10), while a summary of the frequency distribution of underachievers is shown for boys and girls in each school in Table 6.2.

Students whose expected scores matched their postscores, with the deviation being ± 5%, were identified as having matching efficacy-expectancy and are shown in Appendix 11.

On the other hand, for students whose expected scores were less than their postscores, the deviations were negative. Students with negative deviations beyond 5% in magnitude were identified as having lower efficacy-expectancies and are presented in Appendix 12. Table 6.2 reveals that the problem of underachievement was severe across almost all classes because, for instance, in Classes 3, 5 and 12, no student was able to achieve as much as expected. For Class 10, however, the proportion

of underachievers was negligible (with only one of 37 students having underachieved). Possible causes of underachievement are analysed in the next section.

Table 6.2: Number of Underachievers in Each School

School ID	Class ID	No. of Underachievers			Class Size*	Underachievers as % of Class Size
		Boys	Girls	Total		
School 1	1	6	6	12	21	57.1
School 2	3	21	17	38	38	100.0
	4	27	10	37	41	90.2
	5	8	8	16	16	100.0
School 3	6	15	13	28	41	68.3
	7	11	4	15	17	88.2
	8	8	10	18	39	46.2
School 4	9	6	2	18	30	60.0
	10	1	0	1	37	2.7
School 5	11	19	15	34	37	91.9
	12	13	14	27	27	100.0
	13	17	10	27	28	96.0
School 6	14	20	0	20	31	65.0
	15	9	0	9	30	30.0
	16	0	30	30	34	88.2
Sample		191	139	330	464	71.1

Note: * Class size refers to the total number of students who gave responses to the MSES, MCOLES, pretest and the posttest.

6.3.2 Analysis of Efficacy-Expectancies and Causes of Underachievement

To provide an explanation of underachievement, research objective 4(b) of Subsection 1.5.2 was pursued further by investigating the extent to which students' self-efficacy varied with their prior knowledge. Using students' efficacy-expectancy (expected scores) and achievement (postscores), a joint frequency distribution was derived for three categories of students (Table 6.3):

a) underachievers (or students with higher efficacy expectancy) whose expected scores were more than their postscores with the deviations being positive and more than 5%
b) students with matching efficacy expectancy whose expected scores matched their postscores with the deviations being in the range of ± 5%

c) students with lower efficacy expectancies whose expected scores were less than their postscores with the deviations being negative and more than 5% in magnitude.

Table 6.3: Frequency Distribution of Underachievers and Students with Matching or Lower Efficacy-Expectancies

Student Categories	Boys		Girls		All	
	N	%	N	%	N	%
Underachievers	191	67.0	139	77.7	330	71.1
Students with Matching Efficacy Expectancy	39	13.7	15	8.4	54	11.6
Students with Lower Efficacy Expectancy	55	19.3	25	14.0	80	17.2
All Types	285	100.0	179	100.0	464	100.0

Table 6.3 reveals that learning outcomes were not commensurate with expectations. Many students over-assumed their capabilities and could not solve the test tasks successfully. As high as 71% of students in the sample underachieved, accounting for 191 boys and 139 girls or a total 330 out of 464 students in the sample. Interestingly, 54 students (39 boys and 15 girls) or over 11% of the sample had postscores that matched their expected scores with deviations of ± 5%. And, 80 students (55 boys and 25 girls) or over 17 % of students achieved postscores beyond what they expected.

An important aspect of the analysis, which was of prime interest in this study, was to assess to what extent students' prior knowledge, when mediated by self-efficacy, influences their achievement. This was examined for each ability group by using the linkage between self-efficacy and achievement.

Prior knowledge could be a cause of underachievement. Thus, it is important to analyse the joint frequency distribution of underachievers for each ability group. An attempt is made below to identify the proportion of students from each ability group who exhibited high self-efficacy expectancy but low achievement.

Two types of joint frequency distributions of underachievers were generated using students' prescores, expected scores and postscores:

i) The joint distribution of underachievers for each ability group based on prior knowledge (prescores)

ii) The joint distribution of underachievers for each self-efficacy group using expected scores.

Table 6.4 shows the frequency distribution of students segregated across three ability groups (or prior-knowledge groups) and Table 6.5 displays the joint frequency distribution of underachievers for each self-efficacy group.

Table 6.4: Joint Frequency Distribution of Underachievers in Each Ability Group

Ability Group (Prescores)	No. of Underachievers			% of Underachievers		
	Boys	Girls	All	Boys	Girls	All
0 - 40% (Low)	20	18	38	10.5	12.9	11.5
40 - 60% (Med)	50	30	80	26.2	21.6	24.2
60 - 80% (High)	121	91	212	63.4	65.5	64.2
Sample Total	191	139	330	100.0	100.0	100.0

Table 6.4 provides evidence that most students displayed adequate prior knowledge based on their prescores but could not solve the given tasks successfully and therefore underachieved. This accounted for 80 students (50 boys and 30 girls) or over 24% in the sample from the medium-ability group, 212 students (121 boys and 91 girls) or over 64% in the sample from the high-ability group, and only 38 students (20 boys and 18 girls) or 11.5% from the low-ability group. Thus, the number of underachievers with low prior knowledge was relatively small as a proportion in the whole sample.

The composition of the sample according to the categories mentioned in Tables 6.3 and 6.4 is displayed in Figure 6.1. Underachievers, as proportions in the overall sample of 464 students, are shown separately for three ability groups based on prescores. The proportion of all underachievers in the sample was 71.1% (Table 6.3), whereas underachievers from the low-ability group accounted for only 38 out of 464 students or 8.2%; those from the medium-ability group accounted for 17.2%; and underachievers from the high-ability groups were as high as 45.7%. This segregation

suggests that mostly students of the medium- and high-ability groups underachieved, which implies that low prior knowledge was not a major cause of underachievement.

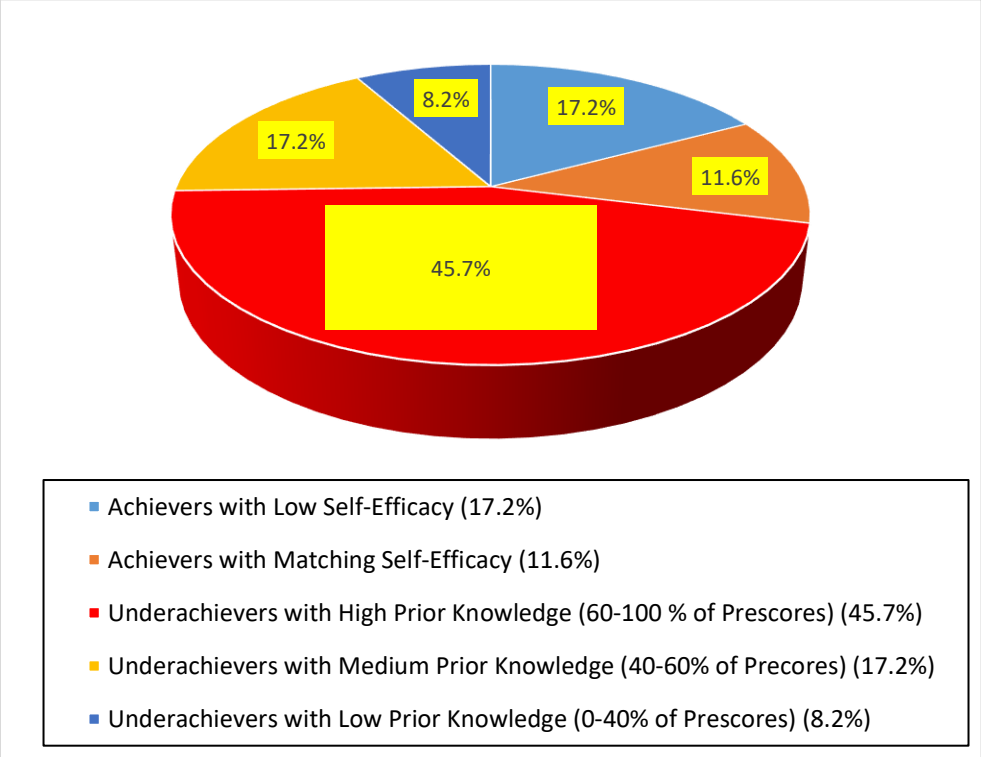


Figure 6.1: Composition of Different Types of Achievers

Table 6.5 further substantiates underachievement for a majority of 80% of boys and 85% of girls from the high-efficacy group, and about 20% of boys and over 11% of girls from the medium-efficacy group, which suggests that a tendency to over-assume and misjudge one's own capabilities to solve give tasks was a major cause of underachievement.

Thus, most students had high self-efficacy based on their prescores, but had over-estimated their achievement, which is a situation of 'high efficacy - low achievement'. It implies that the pretest was rather easy for them. The fact that postscores were lower reflects that students did not master the tasks like those on the posttest. Students with such a tendency might not persevere to solve challenging tasks as much as those students with low self-efficacy (Bandura, 1977). Tables 6.4 and 6.5 together justify the conclusion that many students over-assumed their judgemental capabilities to solve the tasks on the posttest successfully, even though they were assessed as having adequate prior knowledge.

Table 6.5: Joint Frequency Distribution of Underachievers in Each Self-Efficacy Group

Self-Efficacy Group (Expected Scores)	No. of Underachievers			% of Underachievers		
	Boys	Girls	Total	Boys	Girls	Total
0 -40 % (Low)	0	5	5	0.0	3.6	1.5
40-60 % (Med)	38	16	54	19.9	11.5	16.4
60-80 % (High)	153	118	271	80.1	84.9	82.1
Sample Total	191	139	330	100.0	100.0	100.0

Thus, prior knowledge was a cause of underachievement only to a limited extent, while other factors including students' misunderstanding, lack of effort to master such tasks, or/and high efficacy-expectancy with a tendency to over-assume their own capabilities, could be important causes for their underachievement.

Thus, these students might benefit from remedial counselling that involves teachers encouraging them to seek help in or outside the class to improve their understanding of difficult tasks, particularly those types of mathematics tasks for which they scored less than expected.

Lists of all three categories of students are shown in Appendices 10-12, which can be communicated to all schools as a basis for an appropriate teacher intervention program to improve students' mathematics learning in future. The next section considers the predictive validity of the MSES using correlations between efficacy-expectancies and achievement.

6.4 Predictive Validity of MSES Involving Gender Differences in Correlations between Achievement and Efficacy-Expectancies

This section addresses research objectives 4(c) and 4(d) of Subsection 1.5.2 involving the predictive validity of the MSES by an analysis of correlations between achievement and efficacy-expectancies and gender differences in efficacy-expectancies. Subsection 6.4.1 examines bivariate correlations between students' efficacy expectancies and their i) achievement, ii) prior knowledge and iii) classroom

learning environment, while Subsection 6.4.2 reports effect sizes for gender differences in efficacy-expectancies and their statistical significance for different schools and the whole sample.

6.4.1 Predictive Validity of MSES by Analysis of Correlations between Efficacy-Expectancies and Achievement

To investigate the predictive validity of the MSES, bivariate correlations were obtained between efficacy-expectancies and i) achievement, ii) prior knowledge and iii) classroom learning environment. First, bivariate correlations between efficacy-expectancies and achievement are reported in Table 6.6 for boys, girls and the whole sample. Gender differences in correlations were tested using the Fisher's *r*-to-*z* transformation (Lowry, 2017) and results are shown in the last column of Table 6.6.

Table 6.6: Correlations between Self-Efficacy Expectancies and Achievement for Boys and Girls, and Gender Differences in Correlations in Each School

School	Whole Sample		Boys		Girls		Gender Differences in Correlations	
	r_{all}	N_{all}	r_b	N_b	r_g	N_g	Magnitude	Significant?
School 1	0.42	21	0.51	8	0.79	13	-0.29	Yes
School 2	0.48	95	0.36	59	0.05	36	0.31	Yes
School 3	0.23	97	0.72	56	0.20	41	0.52	Yes
School 4	0.57	67	0.38	51	0.57	17	-0.19	No
School 5	0.38	89	0.68	50	0.83	38	-0.15	No
School 6	0.40	95	0.48	61	0.43	34	0.05	No
All	0.42	464	0.43	285	0.39	'9	0.04	No

Note: * Significance was tested by using the Fisher's *r*-to-*z* transformation (Lowry, 2017).

A major finding was that, for the overall sample, the correlations between efficacy-expectancies and achievement were moderate at 0.42 and varied from 0.23 for School 3 to 0.57 for School 4. Correlations varied over a wider range for girls (0.05-0.83) than for boys (0.35-0.72). There were significant gender differences in correlations for Schools 1, 2 and 3, with the highest gender difference in correlations being 0.52 in School 3, where the correlation was higher for boys (0.72) than for girls (0.20). The lowest gender difference in correlation of 0.05 in School 6 was nonsignificant, with the correlation being slightly higher for boys (0.48) than girls

(0.43). Bivariate correlations between efficacy-expectancies and achievement were investigated for each type of task, and for all tasks combined, for each class in all schools, as shown in Table 6.7.

Table 6.7: Bivariate Correlations between Students' Self-Efficacy Expectancies (Expected Scores) and Achievement (Postscores) for Each Class and Task Type on the Posttest.

Class ID	Size (N)	Task Types*						
		D1	D2	P1	P2	P3	P4	All Tasks
1	21	0.75	NS	0.70	0.57	0.69	0.49	0.69
3	38	NS	NS	NS	NS	NS	NS	NS
4	41	NS	0.37	0.32	NS	NS	0.31	0.34
5	16	NS	NS	NS	NS	NS	NS	NS
6	41	NS	NS	NS	NS	NS	0.30	0.35
7	17	0.48	NS	0.57	NS	NS	NS	NS
8	39	NS	0.31	0.36	0.51	0.37	0.33	0.46
9	30	0.46	NS	0.53	0.37	0.47	0.55	0.53
10	37	0.52	0.46	0.43	0.46	0.30	0.52	0.56
11	34	0.42	NS	0.41	0.51	0.43	NS	0.61
12	27	0.36	0.54	0.62	0.78	0.66	NS	0.87
13	28	NS	0.48	NS	0.40	NS	0.47	0.60
14	31	0.51	0.46	0.46	0.42	0.56	0.70	0.64
15	30	0.48	0.40	0.37	0.48	0.48	0.40	0.56
16	34	0.33	NS	0.55	0.43	0.37	NS	0.43

Notes: * Tasks on the posttest were classified into six types.
All estimates above were statistically significant except NS (Not Significant).

In individual classes, students' efficacy-expectancies were moderately correlated with their achievement for each of six types of tasks. The ranges of correlations were: i) 0.30-0.56 in Class 10 of School 4; ii) 0.37-0.70 for Classes 14 and 15 of School 6; iii) 0.49-0.75 for Class 1 of School 1 and iv) 0.36-0.87 for Class 12 of School 5. A major finding was that correlations between students' achievement and their expected scores were relatively stronger when computed for individual classes for specific types of tasks, or all tasks combined, than for all classes together in a school or the whole sample. This finding demonstrates that the calibration method (Bandura, 2006) offers teachers a powerful classroom tool for predicting students' achievement.

The role of efficacy-expectancy was further investigated by analysing its associations with students' prior knowledge and classroom learning environment. For

the overall sample (bottom row of Table 6.8), students' prior knowledge was also found to be moderately and significantly correlated with their efficacy-expectancies, which was comparable with the correlation of 0.35 between prior knowledge and achievement as reported in Table 4.7 of Chapter 4.

Table 6.8: Bivariate Correlations between Self-Efficacy Expectancies and Prior Knowledge for Boys and Girls, and Gender Differences in Correlations in Each School

School	Whole Sample		Boys		Girls		ıder Differences in Correlations	
	r_{all}	N_{all}	r_b	N_b	r_g	N_g	Magnitude r_b - r_g	Significant or not?
School 1	0.45	21	NS	8	0.65	13	-0.65	Yes
School 2	NS	95	NS	59	NS	36	-	NS
School 3	0.39	97	0.55	56	NS	41	0.55	Yes
School 4	0.33	67	0.34	51	NS	17	0.34	Yes
School 5	0.62	89	0.47	50	0.73	38	-0.26	Yes
School 6	0.12	95	0.34	61	NS	34	0.34	Yes
All	0.35	464	0.37	285	0.32	179	0.05	No

Note: * Tested by using the Fisher's *r*-to-*z* transformation and online tool (Lowry, 2017).
All estimates were statistically significant except NS (not significant) at 5%.

The correlation between prior knowledge and self-efficacy was also significant and slightly stronger for boys (0.37) than for girls (0.32), although their difference was not statistically significant at the 5% level. This finding supports a positive relationship between prior knowledge and self-efficacy.

Also, students' efficacy-expectancies as measured by expected scores were correlated with their perceptions of classroom learning environment (Table 6.9). For this purpose, use was made of factor values for the overall classroom learning environment extracted by the second-order CFA of MCOLES, as discussed in Subsection 5.3.4.

The correlation between self-efficacy and classroom learning environment was moderate for boys in Schools 2, 3, 4 and 6, but was nonsignificant for girls in all schools except School 4, where the correlation was stronger for girls at 0.65 than for boys (0.40). For the whole sample, the correlation between efficacy-expectancy and

overall classroom learning environment was low at 0.22. However, this correlation was moderate and higher for boys at 0.31.

Table 6.9: Bivariate Correlations between Self-Efficacy Expectancies and Classroom Learning Environment for Boys and Girls, and Gender Differences in Correlations in Each School

School	Whole Sample		Boys		Girls		Gender Differences	
	r_{all}	N_{all}	r_b	N_b	r_g	N_g	r_b-r_g	Significant?
School 1	NS	21	NS	8	NS	13	-	-
School 2	NS	95	0.36	59	NS	36	0.36	Yes
School 3	0.47	97	0.51	56	NS	41	0.51	Yes
School 4	0.48	67	0.40	51	0.65	17	-0.25	No
School 5	NS	89	NS	50	NS	38	-	-
School 6	NS	95	0.34	61	NS	34	0.34	Yes
All	0.22	464	0.31	285	NS	179	0.31	Yes

Note: * Tested by applying the Fisher's *r*-to-*z* transformation and online tool (Lowry, 2017).
All estimates above were statistically significant except NS (Not Significant) at the 5% level.

Furthermore, considering students' perceptions of Teacher Support, which was identified as a dominant classroom learning environment factor in Chapter 5, correlations with efficacy-expectancies also were calculated, but not shown in any tables. Interestingly, for the overall sample, the correlation between Teacher Support and efficacy-expectancy was also found to be 0.22, which is the same as the correlation for the overall classroom learning environment. For boys, the correlation of efficacy-expectancy with Teacher Support was less at 0.10 than with classroom learning environment (0.31). For girls, efficacy-expectancy was correlated with Teacher Support significantly at 0.18 but was nonsignificant with overall classroom learning environment. These findings support a positive relationship between efficacy-expectancy and classroom learning environment, especially Teacher Support.

6.4.2 Gender Differences in Efficacy-Expectancies

This subsection addresses research objective 4(d) about gender differences in efficacy-expectancies and in correlations between efficacy-expectancies and i) prior knowledge and ii) Teacher Support. First, the means and SDs of expected scores for boys and girls were used for estimating gender differences in efficacy-expectancies

among students in each school, as shown in Table 6.10. Then, for gender differences, effect sizes (*d*) (Cohen, 1988) and statistical significance were estimated.

Table 6.10: Effect Sizes and Significance Tests for Gender Differences in Efficacy-Expectancies for Each School

School	Boys			Girls			Gender Differences	
	N	Mean	SD	N	Mean	SD	Effect Size (*d*)	Significant at 5%?
School 1	8	65.58	22.78	13	53.91	19.29	0.55	No
School 2	59	71.61	16.78	36	68.04	15.14	0.22	No
School 3	56	72.81	17.51	41	73.27	8.80	-0.03	No
School 4	51	64.80	13.67	17	70.42	14.64	-0.40	No
School 5	50	72.43	11.89	38	76.14	12.96	-0.30	No
School 6	61	82.56	12.29	34	77.11	16.54	0.37	No
All	285	72.94	15.89	179	71.91	15.12	0.07	No

Note: Effect sizes were calculated by using the formula given by Cohen (1988).

A key finding was that boys of Schools 1, 2 and 6 had somewhat higher efficacy-expectancies than girls, whereas girls of Schools 3, 4 and 5 displayed somewhat higher efficacy-expectancies than boys. The magnitude of gender differences (effect sizes) were positive and varied from 0.22 to 0.55 for Schools 1, 2 and 6, and were negative (indicating that girls scored better than boys) for Schools 3, 4 and 5 with magnitudes of 0.03-0.40. But gender differences were not significant for all schools. For the overall sample, the effect size for gender difference in efficacy expectancies was small at 0.07 SD, which was not statistically significant at the 5% level.

Earlier, it was reported that students' prior knowledge was positively related with efficacy-expectancies for the overall sample (Table 6.9), which is also corroborated by gender differences in correlations observed for different schools. Correlations between students' prior knowledge and efficacy-expectancies were significant and higher for boys in Schools 3, 4 and 6. In contrast, they were higher for girls in Schools 1 and 5 only, while correlations were almost the same for boys and girls in School 2.

Similarly, gender differences in correlations between efficacy-expectancies and Teacher Support were statistically significant for all individual schools surveyed. For girls, correlations between efficacy-expectancies and Teacher Support were nonsignificant while, for boys, they varied from 0.25 for School 2 to 0.46 for School 3. In Schools 1 and 5, however, correlations were nonsignificant for both boys and girls.

6.5 Chapter Summary

This summary has three subsections. Subsection 6.5.1 summarises results of a) development of the self-efficacy instrument and b) calibration of students' self-efficacy judgements into expected scores that were analysed to identify underachievers, and students with lower efficacy-expectancies. Subsection 6.5.2 summarises key findings about the predictive validity of the MSES that emerged from analyses for gender differences and correlations between efficacy-expectancies and achievement, while Subsection 6.5.3 considers the limitations in calibration and ideas for future search.

6.5.1 Development of MSES and Calibration of Students' Self-Efficacy Judgements

Following a method recommended by Bandura (2006), resource material for the MSES was developed with six types of mathematics tasks. Their content validity was established using a knowledge framework by Hailikari, Nevgi and Lindblom-Ylänne (2007) and a validity framework by Trochim and Donnelley (2006), while adopting the assessment criteria of equity and fairness for designing the resource material for the MSES. Later, the same resource material was used for students' achievement posttest. Face validity was established using feedback from respondents (Munby, 1997). Thus, the MSES was validated and demonstrated its usefulness as a classroom tool.

Furthermore, by calibration of students' self-efficacy judgements, students' responses to MSES items on a scale of 1 to 5 were converted into equivalent expected scores to obtain a measure of efficacy-expectancies.

Efficacy-expectancies measured by expected scores were found to be a valuable source of student data for teachers to assess students' capabilities in achieving. The utility of the MSES as a classroom tool was shown by estimating the proportions in the whole sample of 464 students of the extent to which underachievement was due to prior knowledge or other reasons:

A. underachievement due to inadequate prior knowledge being relatively low (about 8% only)
B. underachievement due to over-assumed capabilities but with reasonable prior knowledge, as shown by students from the medium- and high-ability groups together (about 63%)
C. achieving test scores as expected (about 12 %)
D. achieving test scores beyond what was expected (or students with lower self-efficacy expectancies) (about 17%).

This study, thus, identified students who needed remedial counselling so that their achievement could be improved in the future by an appropriate teacher intervention program. If unchecked, the tendency to over-assume one's own judgemental capabilities might be replicated by students not only in mathematics but also across the curriculum.

A list of students who over-assumed was communicated to teachers in different schools for review. Applying similar procedures, teachers could identify underachievers in their classes and provide remedial counselling for improving students' learning and achievement in the future.

6.5.2 Predictive Validity of MSES and Gender Differences in Efficacy-Expectancies and Correlations

Research objectives 4(c) and 4(d) of Subsection 1.5.2 were examined by investigating the predictive validity of the MSES by an analysis of correlations

between efficacy expectancies and achievement, and gender differences in efficacy expectancies. Students' responses to the MSES were used in calibrating their self-efficacy judgemental capabilities into expected scores for achieving the tasks on the posttest. Students' efficacy-expectancies were moderately correlated with their achievement (postscores). The correlations between expected scores and achievement were stronger for individual classes than in the overall sample. For the overall sample, correlations between expected scores and achievement were lower for girls (0.39) than for boys (0.43), and thus gender differences in correlations were negligible.

For individual schools, however, there were significant gender differences in correlations between efficacy-expectancies and achievement, with correlations being higher for boys in Schools 2 and 3 and for girls in School 1 only. For other schools in the sample, gender differences were not significant.

Interestingly, in individual classes, correlations between efficacy-expectancies and achievement were moderate for six types of tasks: for Class 10 of School 4, they were 0.32-0.56 and were relatively stronger for Classes 14 and 15 of School 6 (0.37-0.70), which established that the validated MSES can be used by teachers to predict students' achievement in their classes. The results of the correlation analysis inform teachers about the relative importance of task types, and where students' learning could be improved in different classes in future. Teachers could use a similar approach for improving their students' learning, particularly after providing remedial counselling to students about exercising appropriate caution in self-reporting their self-efficacy judgements, as recommended earlier.

Correlations between efficacy-expectancies and classroom learning environment were also estimated using expected scores and factor values extracted from second-order CFA of the MCOLES. For the overall sample, correlations were statistically significant, but weak and/or nonsignificant in some schools, suggesting that there were some concerns with students' classroom learning. This finding further supports the usefulness of the MSES and the MCOLES for guiding improvement in students' mathematics learning and achievement.

Regarding the effect sizes for efficacy-expectancies (Cohen, 1988), boys of Schools 1, 2 and 6 had somewhat higher efficacy-expectancies than girls, whereas girls

of Schools 3, 4 and 5 displayed efficacy-expectancies that were marginally higher than those of boys, but their gender differences were nonsignificant in all schools. Gender differences varied between 0.22-0.55 SD in favour of boys in Schools 1, 2 and 6. For Schools 3, 4 and 5, effect sizes in favour of girls were relatively small (0.03-0.40 SD) and, for the overall sample, the effect size for gender differences was small at 0.07 SD but was not statistically significant at the 5% level.

Considering students' perceptions of Teacher Support, which was identified as a dominant classroom learning environment factor, the correlation between Teacher Support and efficacy-expectancy was also estimated at 0.22 for the overall sample, which is also the same as the correlation between self-efficacy and the overall classroom learning environment.

6.5.3 Limitations and Suggestions for Further Research

In calibrating students' judgements, I noticed that a limitation in using the MSES related to the underlying assumption of scale continuity or the gap between any two response opportunities available to students on the scale. When students expected a score between any two consecutive options, the nearest option was only available, which could be higher or lower than what student would have preferred. As a result, it could penalise students incorrectly by classifying them as underachievers. Bandura (2006) recommended that, to minimise such error, scales should be designed with response opportunities as close as possible. This study allowed for ± 5% error.

The second limitation was in using the MSES as a teacher's tool. The MSES needs to be amended in line with the test material designed for any topic of the mathematics curriculum taught by teachers and used for identifying underachievers or those students with lower efficacy expectancies for that topic.

The next chapter reports structural equation modelling of students' achievement including the mediational role of self-efficacy in influencing the effects of prior knowledge and classroom learning environment.

Chapter 7

STRUCTURAL EQUATION MODELLING OF STUDENTS' SELF-EFFICACY AND ACHIEVEMENT

7.1 Introduction

This chapter addresses research objectives 5(a) and 5(b) of Subsection 1.5.2 involving:

a) a mediation analysis of self-efficacy beliefs to estimate direct and mediated achievement effects of students' prior knowledge and classroom learning environment
b) the joint influence of students' prior knowledge and classroom learning environment on achievement by estimating and evaluating a Two-level structural equation model (SEM) of students' achievement.

For research objective 5(a), a mediational model was proposed in Subsection 3.3.2 and, for research objective 5(b), a SEM was described in Subsection 3.3.3 and estimation procedures were discussed in Section 3.7. This chapter presents results of model estimation in six sections. Section 7.2 examines the rationale for adopting Two-level SEM by a variance component analysis when data on dependent variables are hierarchical. Section 7.3 outlines the mediational role of self-efficacy in students' achievement and reports estimates of direct and mediated effects of prior knowledge and classroom learning environment on achievement.

Section 7.4 presents the SEM and its three variants (SEM 1 - SEM 3) using different types of prior knowledge as explanatory variables jointly with classroom learning environment, and then reports maximum likelihood estimates as well as hypothesis testing for the joint influence on achievement.

Section 7.5 provides a comparative evaluation of three variants of the model to identify which one offers the best fit for explaining the variance in students' achievement using model fit indices,

as discussed in Subsection 3.7.4. Finally, Section 7.6 provides a chapter summary with limitations and suggestions for further research.

7.2 Rationale for Two-level SEM by Variance Component Analysis

This section presents an analysis of variance components in two subsections. Subsection 7.2.1 offers a rationale for a Two-level SEM when using hierarchical data, and Subsection 7.2.2 reports the results of variance component analysis.

7.2.1 Rationale for Two-level Structural Equation Modelling

Students' data considered for the model estimation were hierarchical with 464 students drawn from 15 classes from six different schools (five from India and one from Victoria, Australia) and taught by 11 different mathematics teachers at the Year 10 level, with students' achievement varying WITHIN Class and BETWEEN Classes. For selecting a suitable method of estimation, Goldstein (2011) recommends examining the magnitude of intraclass variation (i.e., the proportion of variance in the dependent variable (s) due to variation BETWEEN Classes in total). Multi-level modelling is recommended if the data are hierarchical to the extent that intraclass variation is 10% or above (Goldstein, 2011).

For applying this rule, an attempt is made below to analyse variance in the dependent variables by decomposing the total variance into: WITHIN Class component (Level 1) and BETWEEN Classes component (Level 2). Maximum Likelihood (ML) estimates of variance components were obtained using M*plus* software (Version 7.4) (Muthen & Muthen, 2008) for the whole sample, and separately for the subsamples of boys, girls and each ability group.

7.2.2 Variance Component Analysis

This subsection reports the results of the ML estimates of variance components of students' achievement (ACH) and self-efficacy beliefs in achieving (SEI) using

hierarchical data. Table 7.1 presents the results for the WITHIN Class component, the BETWEEN Classes component and the intraclass variation. The variance in students' achievement (postscores) was estimated at 171.9 for boys, 178.0 for girls, and 180.4 for all students combined. All estimates of variance components of ACH were statistically significant at the 1% level.

The variance in SEI (expected scores) was also seen to be significant but higher than the variance in achievement for all three subsets of samples. For the overall sample, the variance components of SEI were estimated at 185 for boys, 181.6 for girls and 190.5 for all students, which was higher than those of ACH.

For the BETWEEN Classes component, the ML estimates of variances in ACH and SEI were larger than those for the WITHIN Class component. For boys, the variance of BETWEEN Classes component was 384.3, which was more than double the variance observed for the WITHIN Classes part (171.9). Thus, the total variance in ACH was 556.2 (=171.9 + 384.3).

The intraclass variation in ACH, which is the proportion of the variance due to BETWEEN Classes component in total, was 64.1% for the overall sample, 69.2% for boys, and 58.4% for girls. The intraclass variation in SEI was 23.7% for the overall sample and higher for boys (28.5%) than girls (20.1%). Thus, the intraclass variation was high for both the dependent variables, ACH and SEI, which justifies Two-level modelling of the SEM following the rule of thumb recommended by Goldstein (2011).

The M*plus* program also generated ML estimates of grand means of achievement and self-efficacy expectancy for boys, girls and all students combined (Table 7.1). The ML estimates of the grand means of ACH and SEI can be interpreted as average scores obtained by students across all classes and, because they are maximum likelihood estimates, they cannot be treated as 'an average of all class averages' (Goldstein, 2011).

For ACH, the grand mean indicates the average score obtained by students on the posttest across the whole sample, which was slightly more for boys at 55.3% than for girls at 55.1%. Surprisingly, it was much higher at 56.3% for all students combined due to the estimation method, as pointed out above.

Table 7.1: Estimates of Variance Components of Students' Self-Efficacy Expectancy and Achievement

Dependent Variables (ACH, SEI)	Boys (N=285)		Girls (N=179)		All (N=464)	
	ML Ests.#	*p*	ML Ests.	*p*	ML Ests.	*p*
Variance WITHIN Class						
Students' Achievement	171.9	0.000	178.0	0.000	180.4	0.000
Self-Efficacy Expectancy	185.0	0.000	181.6	0.000	190.5	0.000
Grand Means						
Students' Achievement	55.3	0.000	55.1	0.000	56.3	0.000
Self-Efficacy Expectancy	72.2	0.000	70.9	0.000	71.8	0.000
Underachievement (SEI - ACH)	16.9		15.8		15.5	
Variance BETWEEN Classes						
Students' Achievement	384.3	0.000	249.5	0.008	322.8	0.000
Self-Efficacy Expectancy	74.0	0.004	45.8	0.090	59.5	0.001
Intraclass Variation (%)						
Students' Achievement	69.2	-	58.4	-	64.1	-
Self-Efficacy Expectancy	28.5	-	20.1	-	23.7	-

Note: # Maximum likelihood (ML) estimates were obtained by using the M*plus* software.

The ML estimate of the grand mean of SEI suggests that the average postscores expected by boys was at 72.2%, by girls was 70.9% and by all together was 71.8%. Comparing the grand means of SEI and ACH, it can be inferred that, on average, students expected to score much higher than the postscores awarded by their teachers, on average, which mirrors the finding of underachievement, as discussed in Subsection 6.3.2.

Similar results were obtained for each ability group, as shown in Table 7.2. For the WITHIN Classes component, the variation in ACH was found to be statistically significant. It was highest at 265.0 for the low-ability group, followed by the medium-ability group (about 201) and the high-ability group (about 138).

In contrast, the variance in SEI was the lowest for the low-ability group at about 137, followed by the medium-ability group (165.5) and the high-ability group (185.6). An important feature of the high-ability group was that the variance in postscores was lower than the variance in expected scores (SEI). On the contrary, for students of both the low- and medium-ability, the variance in postscores was higher than the variance in expected scores.

Also, for each ability group, a comparison of grand means of ACH and SEI brings out a key aspect of underachievement by boys and girls, separately. It shows that students of high ability displayed higher self-efficacy expectancy (about 75%) than what they achieved (61%) on the posttest, on average. This comparison implies that many high-ability students over-assumed their capabilities to solve mathematics tasks set on the posttest successfully. The extent of under-achievement can be assessed by observing the difference between the average expected scores and average postscores (see the row SEI - ACH in Table 7.2). An important feature was that underachievement was less, on average, for students of the low-ability group (11.8%) than those of the medium-ability group (15.5%) and the high-ability group (13.6%). A pedagogical implication is that teachers might offer remedial counselling for improving students'. understanding and their self-judgemental capabilities in achieving a given set of tasks successfully.

For ACH, the intraclass variation was higher at 70.5% for the high-ability group than for the medium-ability group (59.2%) or the low-ability group (45.3%) but,

Table 7.2: Estimates of Variance Components of Students' Self-Efficacy Expectancy and Achievement for Ability Groups

Dependent Variables (ACH, SEI)	Low-Ability (0-40%)		Medium-Ability (40-60%)		High-Ability (60-100%)	
	ML Ests.#	*p*	ML Ests.	*p*	ML Ests.	*p*
Variance WITHIN Class	N=51		N=110		N=303	
Students' Achievement (%)	265.0	0.009	201.1	0.000	138.5	0.000
Self-Efficacy Expectancy	137.1	0.000	165.5	0.000	185.6	0.000
Grand Means						
Students' Achievement (ACH)(%)	53.1	0.000	52.1	0.000	61.3	0.000
Self-Efficacy Expectancy (SEI)	64.9	0.000	67.6	0.000	74.9	0.000
Underachievement (%) (SEI - ACH)	11.8		15.5		13.6	
Variance BETWEEN Classes						
Students' Achievement, ACH (%)	217.1	0.055	297.8	0.002	328.1	0.000
Self-Efficacy Expectancy, SEI	71.5	0.000	72.3	0.018	49.8	0.014
Intraclass Variation (%)						
Students' Achievement, ACH (%)	45.3	-	59.2	-	70.5	-
Self-Efficacy Expectancy, SEI	34.2	-	30.3	-	20.3	-

Note: # Maximum likelihood (ML) estimates were obtained by using the M*plus* software.

for SEI, the intraclass variation was lower for the high-ability group (20.3 %) than for the other two groups, which was above 30%. This result suggests the desirability of using Two-level modelling for the estimation of the SEM for each ability group in question

7.3 Mediational Role of Self-Efficacy Expectancy: Direct and Mediated Effects

This section presents a mediation analysis in two subsections. First, Subsection 7.3.1 reviews the mediation model that posits theoretical relationships between the explanatory variables, PK and CLE, the mediator variable, SEI, and the criterion variable, ACH. Subsection 7.3.2 reports the results of the hypotheses testing of direct and mediated effects.

7.3.1 Mediation Model

As discussed in Subsection 3.7.2, the mediation model used in the present context was based on the past research (MacKinnon, 2007; MacKinnon, Fairchild, & Fritz, 2007; Raudenbush, & Sampson, 1999). It has the criterion variable, ACH, the mediator variable of self-efficacy expectancy (SEI) and an explanatory variable, X.

The mediation model is represented by the regression equations (7.1) to (7.3) below:

$$ACH = a_1 + c\,X + e_1 \quad \text{----------} \quad (7.1)$$

$$ACH = a_2 + c'\,X + b\,SEI + e_2 \quad \text{----------} \quad (7.2)$$

$$SEI = a_3 + d\,X + e_3 \quad \text{----------} \quad (7.3)$$

X represents, classroom learning environment (CLE) or PK1, PK2 and PK, one by one at a time, where:

- CLE is a latent variable in the model, measured by factor values extracted by second-order CFA of the MCOLES as discussed in Subsection 5.3.4.

- PK1 is students' prior knowledge required for integration of ideas and concepts for solving application tasks of Type P1
- PK2 is students' prior knowledge required for solving challenging tasks of type P2
- PK is students' prior knowledge required to solve all types of tasks combined.

The two variables, PK1 and PK2, carried a total weight of 64% on the posttest, and PK was measured by the aggregate of prescores (%) for all tasks on the pretest. The effects were termed as 'direct' and estimated by simple linear regression of ACH on each explanatory variable shown at equation (7.1). Thus, c is the direct effect of X on ACH, while a_1, a_2 and a_3 are the intercepts, and e_1, e_2 and e_3 are the error terms.

The mediator variable, SEI, was measured by expected scores (%) by calibrating self-efficacy judgements of students' capabilities to solve given mathematics tasks on the posttest using MSES responses, as explained in Subsections 3.6.5 and 6.3.1. SEI represents students' self-efficacy judgements for accomplishing given tasks successfully. SEI therefore mediates the effects of the explanatory variable (X) on ACH. An estimate of c' gives the effect of X on ACH adjusted for mediation (MacKinnon et al., 1997).

The mediated effect can be calculated either by the product method ($b.d$) or by the difference method ($c - c'$). MacKinnon et al. (1995) provided a proof for the algebraic equivalence of these two methods. The estimates of b and d and their standard errors were derived by estimating the regression equations (7.2) and (7.3).

By pairing them with the other explanatory variable, classroom learning environment (CLE), the three regressions of SEM variants are:

SEM1: ACH regressed on PK1 and CLE

SEM2: ACH regressed on PK2 and CLE

SEM3: ACH regressed on PK and CLE.

An important assumption underlying the estimation of the mediation model by the method of maximum likelihood is that the error terms are normally distributed which, in turn, implies that ACH and SEI variables are also normally distributed.

Another assumption is about the direction of causality between SEI and ACH variables. In the mediation model, it was assumed that SEI influences ACH, but Bandura (1977) argued that a reciprocal mediation from ACH to SEI could be possible.

To examine the extent of reciprocal causation (k), I estimated a simple linear regression equation using the whole sample of data:

$$SEI = a + k\ ACH + e \quad \text{-------} \ (7.4)$$

The results of estimation of the mediation model are reported in the following section.

7.3.2 Estimation and Hypothesis Testing of Direct and Mediated Effects

The extent of reciprocal mediation from achievement to self-efficacy expectancy was assessed by the regression equation (7.4). The slope coefficient k was assessed at 0.29 when using the overall sample of 464 students and the boys' sample (N=285), but slightly higher at 0.30 for the girls' sample (N=179). In all three cases, the reciprocal mediation was found to be statistically nonsignificant at the 5% level, which justifies the hypothesised equations (7.1) to (7.3) of the mediation model given in Subsection 7.3.1.

In the context of Two-level modelling, the direct effects (c) were measured by estimating regression equation (7.1) at Level 1 after replacing X by CLE, PK1, PK2 and PK as indicated for the four variants SEM0-SEM3, respectively. Results of regression coefficients (*b*) or standardised coefficients (β) in standard deviation (SD) provide estimates of direct effects.

For estimating the mediated effects, first, regression equations (7.2) and (7.3) were estimated to obtain the values of *b* and *d* coefficients for each of the four variants, SEM0-SEM3. The derivation of mediated effects follows the product method (*b.d*). The results of mediated effects are reported in Table 7.3 for the overall sample, in

Table 7.4 for high-ability students and in Table 7.5 for students of the low- and medium-ability groups combined (who scored less than 60% on the pretest), while the details of derivation of the mediated effects are shown along with respective standard errors and *t*-values in Appendix 13.

The *p*-values associated with β coefficients are shown alongside the coefficients in Tables 7.3 to 7.5 and used for hypothesis testing as follows:

i) A *p*-value ≤ 0.01 indicates that the estimated value of the coefficient is statistically significant at the 1% level.
ii) A *p*-value ≤ 0.05 indicates significance at the 5% level.
iii) A *p*-value > 0.05 indicates that the coefficient is not significant at the 5% level.

Thus, all estimates reported in Tables 7.3 and 7.4 for each of the four explanatory variables were statistically significant at the 5% level when the models were run using the data sets of the overall sample and the high-ability group.

In contrast, for the sample of low- and medium-ability groups combined, the estimates were not significant for the direct and mediated effects of the CLE variable and the mediated effect of the PK2 variable.

To interpret the regression coefficients (*b*) of the explanatory variables, one should consider their units of measurement, whereas the standardised coefficients (β) are independent of units of measurement. Thus, the β coefficient was used to interpret the effect of CLE which was measured by factor values that have no units of measurement.

For example, in Table 7.3, the direct effect of CLE (SEM0) was β = 0.219SD which was significant at the 1% level. It can be interpreted that, assuming other factors remain constant, if students' perceptions of classroom learning environment were better off (or worse off) by 1 SD unit, then their achievement, on average, would improve (or deteriorate) by 0.219 SDs.

The mediated effect of CLE was β = 0.121SD, and also statistically significant. It implies that students' efficacy-expectancy mediated the effect of CLE favourably and improved their achievement by 0.121 SDs.

Because PK was measured by prescore (%), its direct effect on achievement can be interpreted by using b = 0.376 (β = 0.540). It suggests that, if the prior knowledge required for learning the topic was higher (or lower) by 10% (or by 1SD), then achievement would improve (or deteriorate) by 3.76% on average (or by higher achievement of 0.54 SDs), assuming other factors remained constant.

The effect of students' prior knowledge of all tasks combined (PK) on achievement (SEM3) as mediated by self-efficacy expectancies was also statistically significant at b = 0.114 (β = 0.164). This implies that students' efficacy-expectancy mediated the effect of prior knowledge favourably, which contributed to 1.14% higher achievement, on average.

To compare the contributions of prior knowledge of task types PK1 and PK2, direct effects from the model variants SEM1 and SEM2 were considered. The estimates suggest that prior knowledge of both task types (PK1 and PK2) was an important contributor to achievement because the direct effects and mediated effects were positive and statistically significant at the 5% level for the overall sample of 464 students as shown in Table 7.3. Moreover, the ML estimate of prior knowledge of type P1 (PK1) was higher in magnitude at 0.199 (β = 0.383) than that of task type P2 (PK2), b = 0.167 (β = 0.378). A similar inference can be drawn for their mediated effects because of a higher statistically-significant effect for PK1 than PK2 variable (Table 7.3).

Also, results suggested that the direct effects of both variables, CLE and PK, on achievement were higher for the higher-ability group than those for the overall sample (Table 7.3) because, by comparison, the estimates were not only positive and statistically significant, but also were higher in magnitude for the high-ability group than for the overall sample. Similarly, for the high-ability groups, the direct effect of the task type P1 (PK1) was higher than that of the task type P2 (PK2) on achievement.

Also, for the low- and medium-ability groups (Table 7.5), the direct and mediated effects were significant for PK1 and PK variables which were very important contributors to achievement. The contribution of prior knowledge of all tasks combined (PK) (SEM3) was found to be paramount because of its high magnitude (b = 0.503 or β = 0.353). Thus, this finding informs teaching practice that it would be

highly desirable to review and improve students' prior knowledge prior to teaching a mathematics topic. This finding also suggests that, other things remaining constant, if students are helped to improve their prescores, say, by 10%, then achievement could go up by as much as 5%, or the achievement distribution could improve by 0.353 SD for every increase of 1 SD in prior knowledge.

7.4 Two- Level SEM and Estimation of Joint Influence

This section addresses research objective 5(d) of Subsection 1.5.2, which was to quantify the joint influence of prior knowledge and classroom learning environment on achievement in a SEM framework. Subsection 7.4.1 presents a review of the research model as proposed earlier in Figure 1.1 and as explained in Subsection 3.7.3, provides the SEM variants, with the details of measurement of all variables. Results of the model estimation and hypothesis testing of the joint influence of students' prior knowledge and classroom learning environment on achievement are reported in Subsection 7.4.2.

7.4.1 SEM Variants

The structural equation model (SEM) has the criterion variable of students' achievement (ACH) and the mediator variable of self-efficacy expectancy (SEI). It is hypothesised that both students' prior knowledge (X) and classroom learning environment (CLE) directly and positively influence achievement (regression equation (7.5)), and their self-efficacy expectancy (regression equation (7.6)), while these effects are mediated by SEI through its influence on ACH (regression equation (7.7)):

$$ACH = a_1 + b_1\ CLE + c_1\ X + e_1 \quad \text{----------} \quad (7.5)$$

$$SEI = a_2 + b_2\ CLE + c_2\ X + e_2 \quad \text{----------} \quad (7.6)$$

$$ACH = a_3 + d\ SEI + e_3 \quad \text{----------} \quad (7.7)$$

where

Table 7.3: Total Sample: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment from the Mediation Model (N=464)

Model	Explanatory Variables	Direct and Mediated Effects on Students' Achievement (ACH)**					
		Direct Effect			Mediated Effect		
		b	β	*p*	b	β	*p*
SEM0	Classroom Learning Environment (CLE)	6.796	0.219	0.001	3.739	0.121	0.000
SEM1	Prior Knowledge of Type P1, requiring Integration of Ideas, PK1	0.199	0.383	0.000	0.074	0.143	0.000
SEM2	Prior Knowledge of Type P2, involving Application Tasks, PK2	0.167	0.378	0.000	0.052	0.119	0.000
SEM3	Prior Knowledge of all Types of Tasks, PK	0.376	0.540	0.000	0.114	0.164	0.000

Notes: ** Maximum Likelihood Estimates were obtained by using the M*plus* software where *b* represents unstandardised model estimate and β represents standardised estimate.

* The mediated effect was obtained by using the product method suggested by MacKinnon, Fairchild and Fritz (2007).

Table 7.4: High-Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment from the Mediation Model (N=303)

Model	Explanatory Variables	Direct and Mediated Effects on Students' Achievement (ACH)**					
		Direct Effect			Mediated Effect		
		b	β	*p*	*b*	β	*p*
SEM0	Classroom Learning Environment (CLE)	8.414	0.314	0.000	2.792	0.104	0.000
SEM1	Prior Knowledge of Type P1, requiring Integration of Ideas, PK1	0.198	0.353	0.000	0.040	0.071	0.016
SEM2	Prior Knowledge of Type P2, involving Application Tasks, PK2	0.161	0.350	0.000	0.033	0.071	0.000
SEM3	Prior Knowledge of all Types of Tasks, PK	0.422	0.441	0.000	0.088	0.093	0.000

Notes: ** Maximum Likelihood Estimates were obtained by using the M*plus* software where *b* represents unstandardised model estimate and β represents standardised estimate

* The mediated effect was obtained by using the product method suggested by MacKinnon, Fairchild and Fritz (2007).

Table 7.5: Low- and Medium-Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment from the Mediation Model (N= 161)

Model	Explanatory Variables	Direct Effect			Mediated Effect		
		b	β	*p*	*b*	β	*p*
SEM0	Classroom Learning Environment (CLE)	2.468	0.065	0.301	2.496	0.066	0.123
SEM1	Prior Knowledge of Type P1, requiring Integration of Ideas, PK1	0.163	0.156	0.009	0.127	0.121	0.014
SEM2	Prior Knowledge of Type P2, involving Application Tasks, PK2	0.100	0.167	0.039	0.024	0.040	0.553
SEM3	Prior Knowledge of all Types of Tasks, PK	0.503	0.353	0.000	0.187	0.131	0.019

Notes: ** Maximum Likelihood Estimates were obtained by using the M*plus* software where *b* represents unstandardised model estimate and β represents standardised estimate

* The mediated effect was obtained by using the product method suggested by MacKinnon, Fairchild and Fritz (2007).

i) a_1 represents the ACH intercept, a_2 the SEI intercept, and a_3 the intercept of ACH when mediated by SEI from their respective regressions (7.5) - (7.7)

ii) b_1 and c_1 represent the joint influences of classroom learning environment (CLE) and prior knowledge (X) on ACH, respectively

iii) b_2 and c_2 represent the joint influences of CLE and X variables on SEI, respectively

iv) d measures the extent of mediation of SEI on ACH.

It is hypothesised that all effects are positive (Bandura, 1977; Ernest, 1991, 1998; von Glasersfeld, 1985, 2000) and, hence, theoretically all coefficients in equations (7.5) - (7.7) should bear a positive sign.

In the context of Two-level modelling as discussed in Subsection 3.7.3, the above-mentioned regressions provide the estimation of the regressions for the WITHIN Class component, and the estimated intercepts of ACH and SEI (a_1 and a_2) are the latent dependent variables for Level 2 regressions on suitable explanatory variable (s), which could explain the variance in ACH intercept and SEI intercept. The Level 2 regression equations are:

$$\text{ACH intercept} = g_1 + h_1 Z + m_1 \quad \text{----------} \quad (7.8)$$

$$\text{SEI intercept} = g_2 + h_2 Z + m_2 \quad \text{----------} \quad (7.9)$$

where m_1, and m_2 are the error terms; g_1 and g_2 are the grand means (averaged across the whole sample) in ACH intercept and SEI intercept, respectively, while Z is any Level 2 explanatory variable whose value varies BETWEEN Classes but a constant WITHIN Class.

Thus, SEM has three variants SEM1-SEM3:

- SEM1 refers to the model with (CLE, PK1) as the pair of explanatory variables included in the regressions
- SEM2 refers to the model with the pair (CLE, PK2)
- SEM3 refers to the model with the pair (CLE, PK).

ACH and SEI were measured by postscores (%) and expected scores (%) for achievement and efficacy expectancy, respectively, as explained in Subsection 3.6.3,

The construct, classroom learning environment (CLE), was used as a latent explanatory variable which was measured by factor values extracted from the application of second-order CFA to students' responses to the MCOLES as discussed in Subsections 3.6.4 and 5.3.4.

Students' prior knowledge (PK1, PK2 and PK) was measured by the respective prescores (%) on a given set of tasks on the pretest.

For Level 2 explanatory variables, use was made of various alternative variables for BETWEEN Classes, but which are constant WITHIN Class, such as:

- Years of mathematics teaching experience
- Class size
- Class mean prescore (%) of PK
- Class mean prescore (%) of PK1
- Class mean prescore (%) of PK2
- The coefficient of variation (CV) in prescores of PK
- CV in prescores of PK1
- CV in prescores of PK2.

Each of these explanatory variables above was tried by replacing Z in equations at (7.8) and (7.9) and, the one that gave the best fit was considered in the final analysis.

Estimates of joint influences of explanatory variables were obtained from both Level 1 and Level 2 regression equations of SEM1 - SEM3 after replacing Z in equations (7.8) - (7.9) by individual differences in prior knowledge of type P1 (PK1) for SEM1, of type P2 (PK2) for SEM 2, and of all tasks (PK) for SEM 3.

Individual differences in PK1 (or PK2 or PK) were measured by the respective coefficient of variation (CV) in prescores (SD/mean prescore) in each class. The CV ratio was used for comparing the effect of individual differences BETWEEN Classes because the value of SD increases (or decreases) when individual differences increase (or decrease) amongst students WITHIN Class.

The sign of the regression coefficient of CV (or Z) is positive, for instance, in equation (7.8) if the ACH intercept increases (or decreases) whenever the CV value increases (or decreases).

However, the CV increases (or decreases) whenever the SD increases (or decreases) and /or the class mean of PK1 (or PK2 or PK) decreases (or increases) (i.e., whenever the SD increases or decreases more than the mean of prescores, which can only be verified empirically).

7.4.2 Joint Influence of Students' Prior Knowledge and Classroom Learning Environment on Achievement

In this subsection, the models SEM 1 to SEM 3 are estimated to quantify the joint influences of classroom learning environment and students' prior knowledge by using M*plus* software. The results estimated for the regressions at (7.5) - (7.7) are shown in Table 7.6 for the overall sample of 464 students and in Table 7.7 for 303 high-ability students. The estimates of all path coefficients involved were positive as hypothesised on theoretical grounds.

Three major findings emerged for the overall sample. First, considering SEM1, when the influence on achievement (CLE) of PK1 jointly with classroom learning environment was estimated using regressions (7.5), only PK1 was found to be statistically significant at 0.115 (β=0.219). The influence of CLE was not statistically significant, although both PK1 and CLE showed a positive and significant influence on self-efficacy (SEI): β=0.307 for PK1 and β=0.208 for CLE, when estimated from regression equation (7.6). Thus, the joint mediation effects of PK1 and CLE by self-efficacy on achievement were positive and significant because the influence of SEI on ACH estimated from regression equation (7.7) was also positive and significant (β = 0.401).

Second, for SEM2, when the joint influence of PK2 and CLE on achievement was quantified directly from regressions (7.5), only PK2 was statistically significant (β=0.246), but classroom learning environment on achievement (CLE) was not. However, when estimated from regression equation (7.6), both PK2 and CLE were positive and significant in influencing self-efficacy for PK2 (β=0.240) and for CLE (β

= 0.227) and SEI was found to favourably influence ACH (β=0.403) from regression equation (7.7).

Third, for SEM3 variant, when the joint influence of PK and CLE was quantified on achievement directly from regressions (7.5), again only PK was statistically significant (β=0.359), but classroom learning environment on achievement (CLE) was not significant. However, when their joint influence was assessed on self-efficacy from regression equation (7.6), both PK2 and CLE were positive and significant (β=0.392) for PK and for CLE (β=0.192). Moreover, the influence of SEI on ACH estimated from regression equation (7.7) was positive and significant (β=0.362). Thus, the joint mediation effects of PK and CLE by self-efficacy on achievement were positive and significant although their direct effects were not.

Comparing the effects amongst the three SEM variants, the joint influence of prior knowledge and classroom learning environment was greater for SEM 3, when prior knowledge of all types combined (PK) was used together with CLE.

Also, considering the results of the high-ability group (Table 7.7) for variant SEM1, when the joint influence of PK1 and CLE was quantified directly by estimating regression (7.5), both PK1 and CLE were statistically significant (β=0.245) for PK1 and for CLE (β = 0.179). Moreover, when estimated from regression equation (7.6), both PK1 and CLE showed a positive and significant influence on self-efficacy (β=0.161) for PK1 and for CLE (β=0.259). Also, when estimated from regression equation (7.7), the SEI showed a positive and significant influence on ACH (β=0.304) Thus, the joint influence of PK1 and CLE were positive and significant both directly on achievement and when mediated by self-efficacy.

A similar result was observed for the high-ability group of SEM2 when the joint influence of PK2 and CLE was quantified directly by estimating regression (7.5). Both PK2 and CLE were found to be statistically significant: β=0.245 for PK2 and β=0.166 for CLE. Also, when estimated from regression equation (7.7), the SEI showed a positive and significant influence on ACH (β=0.313). Thus, the joint influence of PK2 and CLE was positive and significant both directly on achievement and when mediated by self-efficacy for the high-ability group only.

Another similar result was observed for the high-ability group of SEM3 (Table 7.7) when the joint influence of PK and CLE was quantified directly by estimating regression (7.5). Both PK and CLE were found to be statistically significant at β=0.308 for PK and at β=0.148 for CLE. Also, when estimated from regression equation (7.7), the SEI showed a positive and significant influence on ACH (β=0.284). Thus, the joint influence of PK and CLE were positive and significant both directly on achievement and when mediated by self-efficacy.

Comparatively, for the high-ability group, the SEM3 variant indicated that the joint influence of prior knowledge of all types combined (PK) was paramount because of its significantly larger effect on achievement (0.300) when estimated jointly with classroom learning environment. It is important to note that not only students' perceptions of classroom learning environment had a positive and significant influence on their achievement, but also those effects were mediated positively and significantly by self-efficacy on achievement.

Comparing results obtained for the whole sample of students with those of high ability, the higher the students' prior knowledge, the more effective was classroom learning environment in influencing achievement.

Furthermore, when the direct effects (see Subsection 7.4.1) of prior knowledge and classroom learning environment were compared with joint influences on achievement, the joint influence estimated in the SEM framework was lower, generally, than the corresponding direct effects. For example, results reported for the high-ability group in Tables 7.4 and 7.7, indicate that, while the direct effect of PK on ACH was b = 0.422 (Table 7.4), the corresponding joint effect of PK with CLE was lower at b = 0.253 for the variant SEM3 (Table 7.7), probably due to covariance between PK and CLE considered at Level 1 of the SEM and the coefficient of variation (CV) in PK, which reflects WITHIN Class differences in students' prior knowledge, considered at Level 2 of the SEM.

Similarly, while the direct effect of CLE on ACH was β=0.314 (Table 7.4), the corresponding joint effect of CLE with PK was less at β = 0.148 for SEM3 (Table 7.7). It is not surprising to find that direct effects are biased upwards because of confounding influence of excluded variables (Kendall et al., 1973), as well as the mediational role

of self-efficacy, and thus joint effects are less in magnitude than corresponding direct effects, as well as probably being a better representation of reality assuming other factors remain constant.

For the BETWEEN Classes component of variance, the ML estimates of grand means were obtained from the regressions (7.8) - (7.9), as shown in Table 7.8 for the overall sample of students and in Table 7.9 for high ability students, and were statistically significant and positive for all three models, SEM 1 to SEM 3.

The grand mean represents a 'catch all' effect when other covariables in the regression equation are held constant. Its effect was positive and significant in all models, SEM1 to SEM 3, both for the overall sample of students (Table 7.8) and for high-ability students (Table 7.9). The estimates of grand means were much smaller for achievement (ACH intercept) than self-efficacy expectancy (SEI intercept) for all models on average, which substantiate the extent of underachievement reported earlier in Subsection 6.3.2

Also, these results suggest that individual differences in students' prior knowledge of task types (CV of PK1, or CV of PK2) were not significant, whereas the estimate of the effect of CV of PK was significant for explaining the variance in average achievement (the ACH intercept) in regression (7.8), and the variance in average self-efficacy expectancy BETWEEN Classes (the SEI intercept) in regression (7.9). The following section presents results of model evaluation of SEM 1- SEM 3.

7.5 Evaluation of SEM

This section compares and evaluates SEM 1 to SEM 3 using the regression results for both components at Levels 1 and 2 and the fit indices generated by the M*plus* software program. The R^2 -values of ACH and SEI shown in Table 7.10 reflect the proportion of total variance explained in students' achievement and efficacy expectancies, respectively.

Table 7.6: All Students: ML Estimates of Joint Influence of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as a Mediator from WITHIN Class Part of SEM (N=464)

Model	Explanatory Variables WITHIN Class	Influence on Students' Achievement (ACH)**			Influence on Students' Self-Efficacy Expectancy (SEI)		
		b	β	*p*	*b*	β	*p*
SEM1	Prior Knowledge of Type P1 (PK1) & Classroom Learning Environment (CLE)						
	PK1	0.115	0.219	0.000	0.163	0.307	0.000
	CLE	NS	NS	0.181	6.805	0.208	0.002
	SEI	0.395	0.401	0.000	-	-	-
SEM2	Prior Knowledge of Type P2 (PK2) & Classroom Learning Environment (CLE)						
	PK2	0.109	0.246	0.001	0.108	0.240	0.000
	CLE	NS	NS	0.242	7.368	0.227	0.003
	SEI	0.398	0.403	0.000	-	-	-
SEM3	Prior Knowledge: All Types (PK) & Classroom Learning Environment (CLE)						
	PK	0.250	0.359	0.000	0.271	0.392	0.000
	CLE	NS	NS	0.346	6.471	0.192	0.004
	SEI	0.364	0.362	0.000	-	-	-

Note:** Maximum Likelihood Estimates were obtained by using the M*plus* software, where *b* represents unstandardised model estimates β, standardised estimates. NS not significant at the 5% level.

Table 7.7: High-ability Students: ML Estimates of Joint Influence of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as a Mediator from WITHIN Class Part of SEM (N=303)

Model	Explanatory Variables WITHIN Class	Influence on Students' Achievement (ACH)**			Influence on Students' Self-Efficacy Expectancy (SEI)		
		b	β	*p*	*b*	β	*p*
SEM1	Prior Knowledge of Type P1 (PK1) & Classroom Learning Environment (CLE)						
	PK1	0.140	0.245	0.001	0.103	0.161	0.018
	CLE	4.949	0.179	0.001	8.043	0.259	0.005
	SEI	0.271	0.304	0.000	-	-	-
SEM2	Prior Knowledge of Type P2 (PK2) & Classroom Learning Environment (CLE)						
	PK2	0.114	0.245	0.004	0.075	0.143	0.004
	CLE	4.562	0.166	0.004	8.509	0.276	0.005
	SEI	0.276	0.313	0.000	-	-	-
SEM3	Prior Knowledge: All Types (PK) & Classroom Learning Environment (CLE)						
	PK	0.300	0.308	0.000	0.247	0.227	0.000
	CLE	4.155	0.148	0.007	7.777	0.247	0.000
	SEI	0.253	0.284	0.000	-	-	-

Note: ** Maximum Likelihood Estimates were obtained by using the M*plus* software, where *b* represents unstandardised model estimates β, standardised estimates. NS not significant at the 5% level.

Table 7.8: All Students: ML Estimates of Influence of Individual Differences in Students' Prior Knowledge on Class Average Achievement with Self-Efficacy as a Mediator for BETWEEN Classes Part of the SEM (N=464)

Model	Explanatory Variables BETWEEN Classes	Influence on Class Average Score of Achievement (ACH Intercept)**			Influence on Class Mean Score of Self-Efficacy Expectancy (SEI Intercept)		
		b	β	*p*	*b*	β	*p*
SEM1:	Grand mean	47.75	2.905	0.000	80.119	11.561	0.000
	Individual differences in PK1	NS	NS	0.246	NS	NS	0.084
SEM2:	Grand mean	57.19	3.392	0.000	74.389	10.151	0.000
	Individual differences in PK2	NS	NS	0.924	NS	NS	0.180
SEM3:	Grand mean	41.66	2.535	0.001	52.630	9.842	0.000
	Individual differences in PK	69.79	0.345	0.035	NS	NS	0.880

Note:** Maximum Likelihood Estimates were obtained by using the M*plus* software, where *b* represents unstandardised model estimates β, standardised estimates. NS not significant at the 5% level.

Table 7.9: High-Ability Students: ML Estimates of Influence of Individual Differences in Students' Prior Knowledge on Class Average Achievement with Self-Efficacy as a Mediator for the BETWEEN Classes Part of the SEM (N=303)

Model	Explanatory Variables BETWEEN Classes	Influence on Class Average Post Score (ACH Intercept)**			Influence on Class Average Expected Score (SEI Intercept)		
		b	β	*p*	*b*	β	*p*
SEM1:	Grand mean	48.15	2.901	0.000	79.118	12.183	0.000
	Individual differences in PK1	NS	NS	0.088	NS	NS	0.338
SEM2:	Grand mean	56.15	3.157	0.000	73.442	10.879	0.000
	Individual differences in PK2	NS	NS	0.288	NS	NS	0.421
SEM3:	Grand mean	39.07	2.274	0.004	69.682	10.156	0.001
	Individual differences in PK	109.44	0.517	0.000	NS	NS	0.163

Note:** Maximum Likelihood Estimates were obtained by using the M*plus* software, where *b* represents unstandardised model estimates β, standardised estimates. NS not significant at the 5% level.

For evaluating SEM 1 to SEM 3, the following five statistics were adopted:

i) The coefficient of determination (R^2) and residual variances of dependent variables of regressions of the WITHIN Class component
ii) The coefficient of determination and residual variances of dependent variables of regressions of the BETWEEN Classes component
iii) Log Likelihood (Log L)
iv) Akaike Information Criterion (AIC)
v) Comparative Fit Index (CFI).

For ACH, the highest R^2 value was 40.4% for SEM3, followed by SEM2 (31.1%) and SEM1 (30.9%). For SEI, the highest R^{2-} value was also for SEM 3 (23.9%), followed by SEM1 (17.6%) and SEM2 (13.8%). All estimates of R^2 were statistically significant, though low in magnitude. Thus, SEM 3 was the preferred variant.

A similar conclusion was arrived at by using the residual variance (1 - R^2), which is the proportion of total variance unexplained. It was 1- 0.409 = 0.596 or 59.6% for SEM3, the lowest of all variants. For SEI as well, it was again SEM3 which showed the lowest value of residual variance at 76.1% for the overall sample. Nevertheless, because it was large in absolute terms, it warrants further investigation.

For the BETWEEN Classes part, the regression equations for the overall sample were not statistically significant for the regressions of ACH intercept and the SEI intercept for all three SEM variants (Table 7.10). However, among them, SEM3 is preferable to the other two variants as it yielded the highest R^2 -value of 11.9% for the ACH intercept, but an almost negligible value for the SEI intercept, which is a limitation of this study. Also, SEM3 was better than the other two variants by other criteria including the intraclass which was relatively less for SEM3 at 55.3%, as compared with SEM1 (57.9%) and SEM2 (59.4%).

As discussed in Subsection 3.7.4, the Akaike information criterion (AIC), which offers a measure of loss of information due estimated parameters in a model, was also the

Table 7.10: All Students: Evaluation of Two-Level SEM of Students' Achievement (N=464)

Variance Components	SEM 1 (PK1, CLE)		SEM 2 (PK2, CLE)		SEM 3 (PK, CLE)	
	ACH	SEI	ACH	SEI	ACH	SEI
WITHIN Class (Level 1)						
Total Variance	1.000	1.000	1.000	1.000	1.000	1.000
R^2	0.309	0.176	0.311	0.138	0.404	0.239
p	0.000	0.002	0.000	0.006	0.000	0.000
Residual Variance	0.691	0.824	0.689	0.862	0.596	0.761
p	0.000	0.000	0.000	0.000	0.000	0.000
BETWEEN Classes (Level 2)						
Total Variance	1.000	1.000	1.000	1.000	1.000	1.000
R^2	0.045	0.201	0.000	0.042	0.119	0.001
p	0.562	0.387	0.962	0.502	0.290	0.940
Residual Variance	0.955	0.799	1.000	0.958	0.881	0.999
p	0.000	0.001	0.000	0.000	0.000	0.000

Model Fit Information	SEM 1 (PK1, CLE)		SEM 2 (PK2, CLE)		SEM 3 (PK, CLE)	
Intraclass Variation (%)	57.9	19.1	59.4	21.2	55.3	19.7
LOG L		-3688.2		-3694.4		-3676.3
AIC		7404.7		7415.8		7380.6
CFI		1.0		1.0		1.0

least for SEM3 at 7380.6 units. Similarly, the Log L value, which was negative, was the least for SEM3.

When results were analysed for high-ability students (N=303) (Table 7.11), similar conclusions were reached. For the high-ability group, the regression estimates for the WITHIN Class component suggest that the R^2 values for both ACH and SEI were comparatively lower in comparison with their corresponding results for the

overall sample. However, amongst all variants, SEM3 had the highest R^2 value of 32.1% for ACH and 15.9% for SEI variables.

Table 7.11: High-Ability Students: Evaluation of Two-Level SEM of Students' Achievement (N=303)

Variance Components	SEM 1 (PK1, CLE)		SEM 2 (PK2, CLE)		SEM 3 (PK, CLE)	
	ACH	SEI	ACH	SEI	ACH	SEI
WITHIN Class (Level 1)						
Total Variance	1.000	1.000	1.000	1.000	1.000	1.000
R^2	0.284	0.121	0.277	0.107	0.321	0.159
p	0.000	0.090	0.000	0.081	0.000	0.012
Residual Variance	0.716	0.879	0.723	0.893	0.679	0.841
p	0.000	0.000	0.000	0.000	0.000	0.000
BETWEEN Class (Level 2)						
Total Variance	1.000	1.000	1.000	1.000	1.000	1.000
R^2	0.100	0.167	0.075	0.092	0.268	0.100
p	0.562	0.387	0.593	0.689	0.061	0.509
Residual Variance	0.900	0.933	0.925	0.908	0.732	0.900
p	0.000	0.000	0.000	0.000	0.000	0.000
Model Fit Information	SEM 1 (PK1, CLE)		SEM 2 (PK2, CLE)		SEM 3 (PK, CLE)	
Intraclass Variation (%)	64.8	18.1	68.0	19.3	65.5	19.4
LOG L	-2381.4		-2382.1		-2375.9	
	4790.7		4792.2		4779.8	
CFI	1.0		1.0		1.0	

In contrast, for regressions of the BETWEEN Classes component, the R^2 values were higher for the high-ability group of students than the corresponding regression estimates of the overall sample. For example, for SEM3, the R^2 value was 26.8% for the ACH intercept for the high-ability group, whereas it was only 11.9% for the overall sample. Similarly, the R^2 value of SEI intercept regression was 10.0% for the high-ability group in contrast with a negligible value for the overall sample. But R^2 values were statistically nonsignificant for the whole sample, but significant for the high-ability group. The regression results of the joint influence as indicated by the best model fit (SEM 3) are

presented diagrammatically for the WITHIN Class component in Figures 7.1 and 7.2 for the overall sample and high-ability students, respectively, and the corresponding results for the BETWEEN Classes component are given in Figures 7.3 and 7.4.

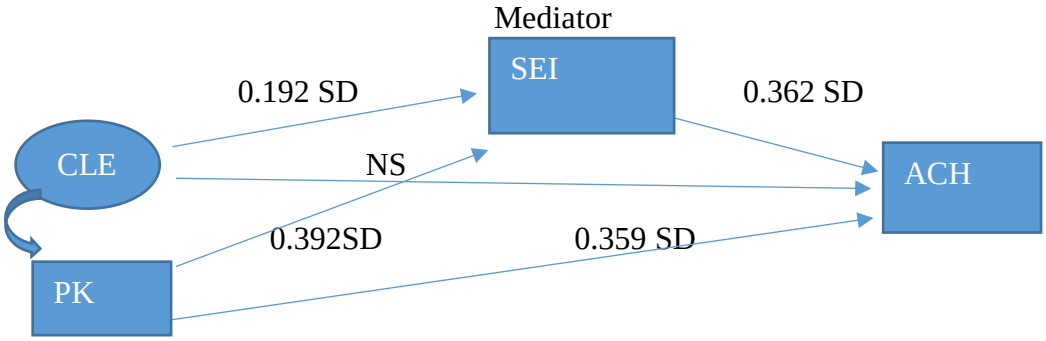
7.6 Chapter Summary

This summary is organised into four subsections that briefly deal with variance component analysis for boys, girls and ability groups (Subsection 7.6.1), direct and mediated effects estimated (Subsection 7.6.2), and ML estimates of the joint influence (Subsection 7.6.3). Finally, Subsection 7.6.4 concludes with limitations and ideas for future research.

7.6.1 Variance Component Analysis

A major finding was that the intraclass variation was much higher at 64.1% for ACH and above 23% for SEI for the overall sample of 464 students. The corresponding estimates were much higher for boys than girls. When students were classified into ability groups according to their performance in a pretest, the intraclass variation in their achievement on the posttest (ACH) was much higher at 70.5% for the high-ability group, 59.2% for the medium-ability group and least at 45.2% for the low-ability group, which implies that there were wide BETWEEN Class differences in achievement within the high-ability group.

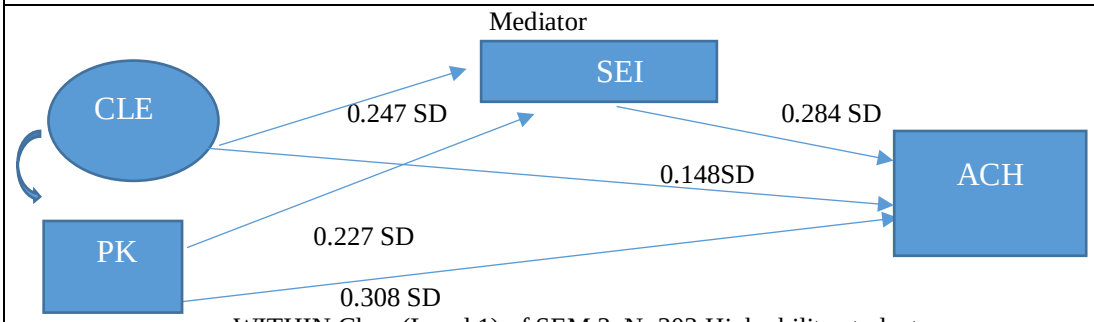
The ML estimates of variances in grand means indicated an important aspect of underachievement by boys and girls, separately. Students of the higher-ability group displayed higher average self-efficacy expectancy (grand mean of SEI) in that their expected scores of 75 % were more than what they achieved (postscores of 61%), on average. This implies that many students of the high-ability group over-assumed their capabilities for successfully solving mathematics tasks set on the posttest, which mirrors a similar finding reported in Subsection 6.3.2. It can therefore be inferred that, by an



WITHIN Class (Level 1) of SEM 3 for N=464 Students

Figure 7.1: All Students: Joint Influence of Classroom Learning Environment and Prior Knowledge of all Types on Students' Achievement as Mediated by Self-Efficacy.

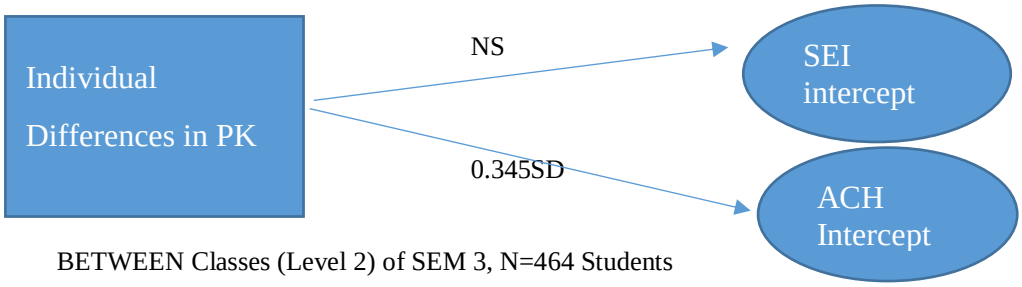
Notes:
1. Prior Knowledge of all Types (PK), Classroom Learning Environment (CLE), Students' Achievement (ACH) and Self-Efficacy Expectancy (SEI) . All observed variables are shown in rectangles and the latent variable, CLE, in ellipse.
2. ML estimates of path coefficients expressed in standard deviation (SD) units obtained by the M*plus* software.
3. All estimates above were statistically significant at the 5% level except that of CLE on ACH (NS= not significant) (see Table 7.6).



WITHIN Class (Level 1) of SEM 3, N=303 High-ability students

Figure 7.2: High-Ability Students: Joint Influence of Classroom Learning Environment and Prior Knowledge of all Types on Students' Achievement as Mediated by Self-Efficacy

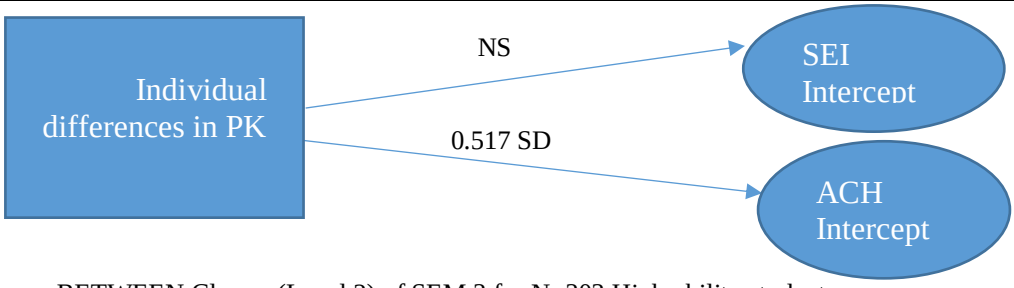
Notes:
1. Prior Knowledge of all Types (PK), Classroom Learning Environment (CLE), Students' Achievement (ACH) and Self-Efficacy (SEI) . All observed variables are shown in rectangles and the latent variable, CLE, in an ellipse
2. ML estimates of path coefficients expressed in standard deviation (SD) units obtained by the M*plus* software.
3. All estimates were statistically significant at the 5% level except those indicated as NS (not significant) (see Table 7.7)



BETWEEN Classes (Level 2) of SEM 3, N=464 Students

Figure 7.3: All Students: Influence of Individual Differences in Prior Knowledge of all Types (CV in PK) on the Average Achievement (ACH Intercept) and the Average Self-Efficacy (SEI Intercept)

Notes:
1. All observed variables are shown in rectangles and the latent variables in an ellipses.
2. ML estimates of path coefficients (Level 2) expressed in SD units were obtained by the M*plus* software (7.4)
3. All estimates were statistically significant at the 5% level except those indicated NS (not significant) as shown in Table 7.8,



BETWEEN Classes (Level 2) of SEM 3 for N=303 High-ability students

---

Figure 7.4: High-Ability Influence of Individual Differences in Prior Knowledge of all Types (CV in PK) on the Average Achievement (ACH Intercept) and the Average Self-Efficacy (SEI Intercept)

Notes:
1. All observed variables are shown in rectangles and the latent variables in ellipses.
2. ML estimates of path coefficients (Level 2) expressed in SD units were obtained by the M*plus* software (7.4)
3. All estimates were statistically significant at the 5% level except those indicated NS (not significant) as shown in Table 7.9.

appropriate teacher intervention, students generally can be offered help and remedial counselling to verify their (mis) understanding and improve their learning.

7.6.2 Direct and Mediated Effects on Students' Achievement

The direct effects of prior knowledge and classroom learning environment derived from the mediation model revealed that classroom learning environment as perceived by high-ability students had a greater direct effect (β=0.314) on achievement, than that perceived by low- and medium-ability groups together (β=0.065), which was not statistically significant at the 5% level. For the overall sample, the mediated effect of classroom learning environment was less on achievement at 0.121SD units than the corresponding direct effect (β=0.219SD).

In contrast, the direct effect of prior knowledge of all types of tasks combined was statistically significant and higher for the low- and medium-ability groups combined (b = 0.503 or β = 0.353) than the high-ability group (b = 0.422 or β = 0.441) which suggests that, while other classroom learning factors remain constant, improving prior knowledge of students from the low- and medium-ability groups by teachers' special efforts, so that students' prescores are 10% higher, would enhance achievement by 5% on their posttest, or improve the achievement distribution by 0.353 SDs. For the overall sample, the mediated effect of prior knowledge was smaller than the respective direct effect.

A comparison of the direct effect of prior knowledge of task types P1 (PK1) with those of type P2 (PK2) revealed that the former type had greater impact on achievement than the latter. Also, results for the higher-ability group (N=303) (Tables 7.3 and 7.4) suggest that the direct effects on achievement of both variables, CLE and PK, were higher for the high-ability group than for the overall sample because, by comparison, the estimates were higher in magnitude for the high-ability group than for the overall sample.

7.6.3 Joint Influence of Classroom Learning Environment and Prior Knowledge on Self-Efficacy and Students' Achievement

To address research objective 5 (b) of Subsection 1.5.2, the research model proposed earlier in Figure 1.1 and explained in Subsection 3.7.3 was specified in three variants, SEM1-SEM3:

- SEM1 posits the joint influence of classroom learning environment (CLE) and students' procedural knowledge for solving tasks of type P1(PK1) on students' achievement (ACH) through the mediator, self-efficacy expectancy (SEI)
- SEM2 models the joint influence of CLE and students' procedural knowledge of solving mathematics tasks of type P2 (PK2) on ACH through the mediator, SEI
- SEM3 hypothesises the joint influence of CLE and students' prior knowledge of solving mathematics tasks all types (PK) on ACH through the mediator, SEI.

The results of hypothesis testing of joint influences were compared between variants and evaluated using the statistics such as the coefficient of determination (R^2), residual variance, Log L and the model fit information including the Comparative Fit Index (CFI) and the Akaike Information Criterion (AIC). Based on detailed evaluation, model SEM 3 was found to be the best representation of the joint influence of (CLE, PK) on students' achievement (ACH).

A major finding was that, for a mathematics topic of measurement that teachers taught to high-ability Year 10 students, the joint influence of classroom learning environment (CLE) and prior knowledge (PK) on achievement was favourable and significant, with the effect of PK being $\beta = 0.308$SD and that of CLE being $\beta = 0.148$ SD. For the whole sample, however, when estimated jointly from the SEM, only the effect of PK was statistically significant, but that of CLE was not.

Comparing the effects amongst the three variants, the joint influence of prior knowledge and classroom learning environment was greater for SEM 3, when prior knowledge of all types combined (PK) was used together with CLE, but the influence of the latter was not statistically significant in any SEM variant, although their effects were mediated significantly and almost equally by self-efficacy on achievement.

Thus, classroom learning environment had a nonsignificant effect on achievement when students did not have adequate prior knowledge required for learning, which is plausible because prior knowledge was found to covary with classroom learning environment and had a bivariate correlation of 0.35 with achievement, as noted in Subsection 5.4.1.

A comparison of estimates of joint influence of CLE and PK indicated that high-ability students perceived their classroom learning environment as important and significant for improving achievement to the extent of 0.148 SDs rather than all students in the sample.

Students' prior knowledge required for learning influenced achievement favourably and statistically significantly (β = 0.359 for the whole sample and β = 0.308 for the high-ability group). This was relatively less than the corresponding direct effect of β = 0.540 for the overall sample and β = 0.441 for the high-ability group, probably due to covariance between PK and CLE.

The influence of CLE, obtained jointly with prior knowledge and as mediated by self-efficacy expectancy, was also statistically significant and greater in magnitude (β=0.247) (Table 7.7) for the high-ability group than for the overall sample (β = 0.192) (Table 7.6).

In general, the estimates of joint influence tended to be smaller in magnitude than their corresponding estimates of direct and mediated effects from the mediation model, which can be attributed to possible covariances between the explanatory variables included in the SEM. The estimates of joint influence thus offer a closer approximation to reality in the assessment of the impact that students' prior knowledge and classroom learning environment together had on their achievement.

7.6.4 Limitations and Suggestions for Future Research

First limitation of this model was that, because the variance explained in the criterion variable (ACH) was almost 40% in the WITHIN Class component and almost

12% only in the BETWEEN Classes component, there was still a substantial degree of unexplained variance in random intercepts in ACH and SEI in the model.

The second limitation was that no estimates of direct and mediated effects could be generated separately for the low- and medium-ability groups because the sample size of 51 low-ability students from 15 classes was inadequate to carry out the estimation of 14 parameters involved in the maximum likelihood estimation of Two-level SEM.

Similarly, the joint influence could not be assessed separately for boys and girls, or for low- and medium-ability groups, because of sample size of 15 classes in the sample of 464 students restricted the researcher to include only one explanatory variable at a time in Level 2 regressions. The explanatory power of the SEM could be improved with a larger sample and more explanatory variables as suggested below.

Future research should involve a larger sample of classes and other explanatory variables (e.g. student cooperation, peer tutoring, parental influence, homework, clubs) to enhance the explanatory power of the Two-level SEM. For higher-level modelling, incorporating students' socio-economic status as a Level 3 variable might be useful to explain variance in students' achievement and self-efficacy BETWEEN schools. This might include investigation by random slopes models in addition to random intercepts models.

Chapter 8

DISCUSSION AND CONCLUSION

8.1 Introduction

In this chapter, I discuss the findings, identify contributions and limitations of my research and offer suggestions for future research. Section 8.2 discusses findings detailed in Chapters 4 to 7 and, where possible, makes links to the past research reviewed in Chapter 2, with possible implications for teaching practice. Section 8.3 summarises the theoretical, methodological and practical contributions of the study. Section 8.4 provides limitations of the study and Section 8.5 offers suggestions for future research. I conclude this thesis with a final comment in Section 8.6.

8.2 Discussion of Findings

The discussion of findings is presented in Subsections 8.2.1 to 8.2.5 according to the order of five research questions of Subsection 1.5.1, along with a brief comparison of results with the past research reviewed in Chapter 2, as well as possible implications for teaching practice.

8.2.1 How Teachers Used Their Students' Prior Knowledge for Classroom Teaching

To examine the first research question, I consulted the 11 mathematics teachers of Year 10 who participated in this study, selected a mathematics topic of measurement and used prescores of a common pretest of students' prior knowledge in that topic to divide their students into three ability groups. Then, they selected the topic of measurement from

the Year 10 mathematics curriculum and taught it over a period of two to three weeks using supplementary tasks suited to students' levels of ability and recorded their classroom teaching experiences to address the four questions in the template of Teachers' Reflective Journal (TRJ).

To the first question of TRJ, which was about the learning intentions for the new topic to be taught, all teachers replied that they adopted the method of whole-class teaching for introducing the learning intention of the topic of measurement by showing various 3D objects of daily life that represented different geometrical shapes. These responses identified an effective way of communication to a large mixed-ability group of students while motivating them for learning the relevant formulae of perimeter, total surface area and volume of basic shapes. This finding was also consistent with the constructivist method of teaching in which teachers verify their students' prior knowledge before starting to teach any new topic and use student assessment data and prior learning to set learning goals. Hattie (2009) found that this teaching strategy has a high impact on learning with an effect size of 0.56 for setting goals.

The second question was about how teachers taught students in the pre-classified ability groups. Many teachers replied that they gave prompts to motivate low-ability students, assigned them open-ended tasks that were challenging, and reinforced their understanding with more worked examples and practice in the classroom and homework on similar tasks. This was consistent with the high impacting teaching strategy of using worked examples, which is a demonstration of the steps required to complete a task or solve a problem. By scaffolding the learning, worked examples support skill acquisition and reduce the cognitive load for learners (State Government of Victoria, 2017). Hattie (2009) found an effect size of 0.57 for the teaching strategy that uses worked examples.

The third question of the TRJ identified how teachers encouraged students of each ability group to facilitate learning. Teachers met low-ability students on a 1-1 basis when they were too shy to ask for help publicly, and often offered them modified work involving visual-learning tasks (such as working out the perimeter and area of basic shapes with the aid of physical objects) or tasks that suited their family background. After inspiring them with encouraging words like '*you can do it'*, teachers allowed them to work collaboratively

with their peers in the classroom, or use visual-learning aids, as in differentiated teaching, which is another high impacting teaching strategy that teachers use to extend the knowledge and skills of every student in every class, regardless of their starting point (State Government of Victoria, 2017).

The importance of activating students' prior knowledge was also supported by the corresponding direct effects obtained in my study using the hierarchical data of 464 boys and girls from the mediational model with the regression coefficient being, b = 0.503, which was significantly higher for low- and medium ability groups than b=0.422 for the high-ability group (for more details see Subsection 7.3.2). This suggests that, when other classroom learning factors remain constant, if the prior knowledge of students from the low- and medium-ability groups combined is improved by teachers' special efforts so that students' prescores are 10% higher, then it would improve students' learning by 5% on their posttest, or the achievement distribution would improve by $\beta = 0.353$ SD. Hattie found an effect size of 1.07 for such response to teaching intervention, as reported in State Government of Victoria (2017).

In summary, the teachers' feedback revealed that many teachers surveyed made effective use of students' prior knowledge in introducing the new topic, by targeted teaching of each ability group using suitable tasks to cater for their cognitive needs in mixed-ability classrooms without adopting problematic tracking or streaming practices (Askew, 2015; Sullivan, 2015).

8.2.2 Gender Differences in and Correlations between Students' Prior Knowledge and Achievement

To address research question 2(a) (Subsection 1.5.1), students' scores in the pretest and posttest were used to obtain gender differences in mean scores and their effect sizes, using the Cohen's (1988) formula, and *t*-tests were used to examine gender differences in prior knowledge and achievement. The first major finding was that, for the topic taught at Year 10 and assessed by teachers, gender difference in students' prior knowledge was not statistically significant at the 5% level, with the effect size (*d)* being only 0.05 SD for the

overall sample of 464 students (285 boys and 179 girls), although the average score of boys was a little better at 67.30% than that of girls at 66.20%.

In Schools 1-3, gender differences in prior knowledge were statistically significant in favour of boys, with the effect size being the highest for School 1 at 0.95 SD, followed by School 3 (0.76SD) and School 2 (0.39SD) whereas, in Schools 4 - 6, girls scored significantly better than boys.

Also, there were wide gaps in prior knowledge amongst students, as reflected by the coefficient of variation (CV ratio) in their prescores, which was 0.31 for boys and 0.33 for girls. In School 6, however, the CV values were small at 0.05 for girls and 0.20 for boys who were disciplined learners. This finding is a distinct contribution because no comparable studies, as reviewed in Chapter 2, investigated gender differences in prior knowledge. These findings suggest that considerable challenges were faced by teachers in catering for the individual cognitive needs of students of all levels of ability in mixed-ability classrooms. To grapple with these challenges, teachers in my research divided students into ability groups within their classes based on students' prior knowledge and taught as discussed in Subsection 8.2.1.

For a topic of measurement selected at the Year 10 level, my study revealed that gender differences in achievement were significant, with the mean postscore of boys being higher at 59.4% than that of girls (55.8%) on average and with the effect size being 0.17 SD.

In Schools 5 and 6, gender differences in postscores were statistically significant, with girls performing better than boys in School 5 (effect size of -0.36 SD) but boys performing significantly better than girls in School 6 (very large effect size of 2.11 SD). But, in Schools 1-4, gender differences were nonsignificant although, for School 4, the girls' mean was higher at 77.50% than that of boys at 71.50%, with an effect size of 0.35 SD.

For India, the NCERT study conducted by Sreekanth et al. (2015) provides evidence of no significant gender differences in mathematics achievement at the Year 10 level. For other countries, however, Else-Quest et al.'s (2010) meta-analytical study using 2003 PISA data reported that effect sizes for gender differences in mathematics achievement across 69

nations were small (0.15 SD), but that the national effect size for Australia, based on participation of 6171 girls and 6380 boys at the Year 10 level, was smaller at 0.06 SD only. Similarly, the weighted mean effect size for gender differences of TIMSS data for the content domains of algebra, data, geometry and measurement and number was almost negligible (0.01 SD) (Else-Quest et al., 2010).

In summary, for the whole sample, my findings did not support gender differences in prior knowledge, while they did for mathematics achievement. Also, for individual schools, gender differences in postscores were statistically significant, with boys scoring more highly than girls in some schools, and girls scoring more highly than boys in some other schools. These results are consistent with the findings based on a public survey in Australia that "many students in this group do not gender stereotype mathematics or English" (Forgasz & Leder, 2015, p. 103).

My research question 2(b) (Subsection 1.5.1) involved an analysis of correlations between students' achievement and i) prior knowledge required for learning a mathematics topic, and covariables including ii) teaching experience and iii) class size. A major finding was that the bivariate correlation between students' achievement and prior knowledge was moderate at 0.35, which was statistically significant for the whole sample.

Considering different types of tasks, the main finding was that correlations between achievement and prior knowledge varied from a low of 0.23 for task type D1 (Declarative knowledge of concepts and meanings) to 0.24 for D2 (Declarative knowledge to recall facts) to 0.32 for P1 (Procedural knowledge of integration of concepts, meanings and facts) and to 0.30 for P4 (Procedural knowledge of applications of problem solving tasks). The correlation was relatively high for tasks of type P1, which is consistent with the fact that, in influencing students' achievement, prior knowledge of solving application tasks (Type P1) assumes greater importance than other types, and therefore teachers must emphasise on this aspect in classroom teaching, while activating prior knowledge using problem-solving tasks.

The bivariate correlation between students' achievement and teaching experience was moderate at 0.31 when the class was used as a unit of analysis. Further, the correlation between students' achievement and class size was relatively high and significant (0.37).

This being positive, contradicts the usual expectation that a teacher's time spent with each student could be higher in classes of smaller size and, hence, students' achievement is likely to be higher for smaller class size, if other factors have remained unchanged.

But it is argued that the correlation between class size and achievement could be positive, possibly because, as a teaching strategy, teachers in the sample divided each class into three ability groups of smaller size, and targeted ability groups by teaching them with enriching tasks suitable to their cognitive needs and levels of ability. Thus, sometimes, students in large classes could even score higher on the posttest than those taught in small classes, if large classes were better organised than the smaller ones.

Apparently, this effort could be confounded by co-covariables (e.g. a large class size, teaching experience and prior knowledge). Also, there was a measurement error for the class size variable, with the number of students attending being higher than that considered in measuring class size, etc. (see limitations as discussed later in Subsection 8.4), which impacted the mean values of prescores and postscores.

Obviously, the larger the class size, the greater the teaching challenge that calls for greater teaching effort for engaging all children cognitively. For example, as discussed in Subsection 4.5.3, in Class 10 which had 41 students, the postscore mean was much higher at 83.9% than the prescore mean of 51.1% and, for Class 15 which had 30 students, the respective mean score was also high at 83.7%, which suggests that ability grouping proved useful for teachers to redress an otherwise negative influence of class size on achievement.

My research objective 2(c) involved examining teachers' selection of tasks to examine whether they were suitable for improving learning and achievement (postscores) of students in mixed-ability classrooms. This was inferred from frequency distributions of prescores and postscores attained by students of each ability group.

For examining changes in learning, when postscore was compared with prescore for each student, only 31.7% or 147 out of 464 students in the sample achieved higher scores in the posttest, which comprised 60.8% or 31 out of 51 students of low ability, 52.7% or 58 out of 110 students of medium ability and 19.1% or 58 out of 303 students of high ability. This finding supports the teaching strategy adopted by participating teachers in my

study that involved a judicious selection of mathematics tasks suitable for ability groups. To a certain extent, teaching strategy addressed the cognitive needs of students, weighing more in favour of students of low and medium ability rather than of higher ability. In the high-ability group, however, a large proportion of 80.8% (245 out of 303 students) did not improve when compared with their prior knowledge.

In School 4, there was a significant improvement in learning among boys, with their mean score increasing from 60% on the pretest to 71.5% in the posttest, with the effect size being 0.90 SD. In Schools 1, 4 and 5, both boys and girls showed improvement in learning, while boys in School 6 and girls in School 3 showed a small decline in learning, whereas girls in School 6, and both boys and girls in School 2, displayed a significant decline in learning.

These results can be claimed to be conservative in that the assessment tasks included in the pretest were relatively easier than those of the posttest. However, the question of how difficult the assessment tasks included in the latter were depends on the levels of mathematical abilities of students at the time of pre- and posttests. Any improvement in their abilities can be attributed to their learning growth or in part the teaching effectiveness, and thus needs further investigation, possibly by employing an alternative educational measurement of assessment under the item response theory (IRT) (Wu, Tsung, & Jen, 2016).

To further address research objective 2(c) of Subsection 1.5.2, results of students' achievement are elaborated for all three ability groups in Classes 1 and 10 that displayed improvement as against Classes 3 and 7 which experienced a decline in learning. For this, use was made of frequency distributions of test scores with means and coefficients of variation (CV).

In Class 1, five out of nine students in the low-ability group and five out of seven from the medium-ability group moved into the higher range of postscores of achievement test. Consequently, the mean score of this class improved significantly to 57.6% from a low of 46.1% on the pretest.

In Class 10, most students achieved higher scores of 60-100% in the posttest, which comprised 6 out of 7 students in the low-ability group and 15 out of 18 in the

medium-ability group. Also, all 12 students of the high-ability group continued to score between 60-100%. Thus, all 37 students of Class 10 contributed to a steep increase in their mean postscore to 83.9 % from an average prescore of 51.1%, which also narrowed down inter-student differences as reflected by a decrease in the CV value from 0.30 for prescores to 0.17 for postscores.

Thus, in both Classes 1 and 10, the correlations between prior knowledge and achievement were significant for each of the four types of tasks (Hailikari, Nevgi, & Lindblom-Ylänne, 2007) despite large class sizes (21 students in Class 1 and 37 in Class 10), with the mean postscores having increased in Class 1 to 57.6%, despite a high CV value of 0.43 in the prescore. The disadvantage of large class sizes was internalised as teachers in these two classes divided their students into smaller groups based on prescores, which probably helped the teachers to focus on teaching students according to their levels of ability, and it paid off with a remarkable improvement in their achievement. These findings in my study strongly support the usefulness of the strategy of ability-grouping without tracking or streaming (Askew. 2015; Sullivan, 2015).

In contrast, in Class 7, there was a decline in learning despite a small class size of 18 students (six in the medium-ability group and 11 in the high-ability group). The mean postscores of the whole class declined because eight students scored less than 40% and the remaining nine students scored between 40-60% only. Also, the gap in performance widened between them, as the respective CV value increased to 0.34 in the posttest from 0.24 in the pretest, which needs further investigation.

8.2.3 Validation of MCOLES, Its Associations with Achievement and Gender Differences in Associations

My research objective 3(a) (Subsection 1.5.2) was to develop and validate a theoretically-sound instrument to measure students' perceptions of classroom learning environment for Year 10 mathematics. For this, I developed the new Mathematics-related Constructivist Oriented Learning Environment Survey (MCOLES) involving a two-stage approach guided by Trochim and Donnelly's (2006) framework with an extensive review of relevant literature. These supported a more-concise 50-item MCOLES with seven

dimensions for mathematics classes after modifying a pre-validated 11-dimensional 88-item COLES. This improved the reliability for each scale and made the MCOLES more concise with a smaller number of 50 items and requiring less response time.

The discriminant validity of the MCOLES was established by examining partial correlations between scale means instead of bivariate correlations to minimise the confounding influence of items. The mean partial correlations between scales were very low (0.08-0.16) as they should be theoretically.

The concurrent validity of MCOLES was established using the eta^2 statistic from one-way ANOVA between 15 classes (independent variable). Eta^2 was statistically significant for all scales (0.08-0.35), thereby supporting the validity of the MCOLES. The study has thus validated the MCOLES with improved scale reliability not only for the whole sample of 511 students, but also for the boys' subgroup (N=313) and the girls' subgroup (N=198) separately.

In the next step, research objective 3(b) of Subsection 1.5.2 was accomplished with EFA applied to each of seven dimensions of MCOLES, one by one. The application of EFA was justified with a meaningful interpretation of items sub-grouped in the process of validation, supported by scree plots for the whole sample of 511 students and the boys' group of 313 observations. However, only nine factors were extracted for the girls' group of 198 observations because Student Cohesiveness and Personal Relevance loaded on one factor only.

Then, to account for the possible variance-covariance structure of students in co-educational classrooms, CFA was applied to MCOLES responses of the whole sample of 511 students simultaneously on all seven dimensions, using the M*plus* software (Version 7.4). It resulted in a ten-factor simultaneous solution with all the maximum likelihood estimates being statistically significant. The model fit was also satisfactory and the values of R^2 for items that loaded on each of ten factors were also statistically significant. These results validated all measurement properties of the MCOLES with its variance-covariance structure of ten factors that best fits students' perceptions of classroom learning environment in mathematics.

A careful review of these ten factors indicated that they were cross correlated with a moderate to strong interrelationships supported by a theoretical rationale, except for Task Orientation and Peer Differentiation for which the correlations were weak (0.04-0.24). When the first-order factors were cross-correlated, second-order CFA using M*plus* software yielded a single-factor solution with factor values of the classroom learning environment (CLE). A single factor representation of CLE was very useful for the predictive analysis of the MCOLES and structural equation modelling of students' achievement.

My research objective 3(c) of Subsection 1.5.2 was accomplished when predictive validity was tested using correlations between students' mathematics achievement and each MCOLES factor. For the whole sample, the correlation of the overall classroom learning environment with achievement was statistically significant at 0.28. For individual MCOLES factors, the correlations of achievement with perceived Teacher Support was 0.28, followed by Equity (0.23), Clarity of Assessment Criteria & Feedback (0.21) and Cooperation (0.21), but was weak with other factors, with Personal Relevance having the lowest correlation of 0.10 with achievement. The correlation between students' prior knowledge and achievement was relatively stronger at 0.35 than with any MCOLES factor.

The final version of the survey has high content, face, convergent, discriminant, concurrent and predictive validities when used in secondary mathematics classes. These results suggest that this survey is likely to be valid and reliable as a classroom tool to guide teachers in improving their teaching practice.

The next research objective 3(d) of Subsection 1.5.2 involved gender differences in correlations between students' perceptions of classroom learning environment and achievement. When bivariate correlations between students' achievement and classroom learning environment were estimated for the whole sample taking all MCOLES factors combined, there were no significant gender differences.

For the boys' group, correlations between achievement and eight of MCOLES factors varied between 0.33 and 0.54 and were significant and stronger for Student

Cohesiveness followed by Equity, Task Differentiation, and Clarity of Assessment Criteria & Feedback, but were nonsignificant for Task Orientation and Peer Differentiation.

In contrast, for the girls' group, the correlation between achievement and Personal Relevance was stronger than for other scales including Involvement and Co-operation. Gender differences in correlations were statistically significant for Student Cohesiveness and Personal Relevance, but nonsignificant for the other eight factors.

In Schools 2 and 6, the correlation between classroom learning environment and achievement was stronger for boys than girls whereas, in School 4, the correlation of CLE with achievement was stronger for girls than boys. In School 3, the bivariate correlation between all factors (combined) of classroom learning environment and achievement was moderate for boys at 0.47 and for girls at 0.35, but the gender difference in these correlations was statistically nonsignificant.

However, when partial correlations were examined between achievement and all MCOLES factors, gender differences were significant. For example, between Teacher Support and achievement, the partial correlation was higher for boys (0.30) than girls (0.15). For other MCOLES factors, gender differences were significant because partial correlations were significant only for boys, but not for girls.

8.2.4 Development of MSES and Calibration of Students' Self-Efficacy Judgements

My research objective 4(a) was to develop the new Mathematics Self-Efficacy Scale (MSES) following a method recommended by Bandura (2006). To accomplish this, I developed resource material for the MSES with six types of mathematics tasks, followed by the response format for the MSES. The content validity of resource material was established using a knowledge framework by Hailikari, Nevgi, and Lindblom-Ylänne (2007) and a validity framework by Trochim and Donnelley (2006), while adopting the assessment criteria of equity and fairness. This resource material was also used for the posttest of students' achievement. Thus, its face validity was established by using feedback from teachers who used similar tasks in their classes for maintaining the confidentiality of the test material.

Furthermore, my research objective 4(b) (Subsection 1.5.2) required the calibration of students' self-efficacy judgements, for which students' responses to the MSES items on a scale of 1 to 5 were converted into expected scores as a measure of efficacy-expectancies. They were found to be a valuable source of student data for teachers for assessing students' capabilities in successfully achieving mathematics tasks on a test. The utility of the MSES as a classroom tool was shown by the extent to which underachievement was due to prior knowledge, inferred from the sample proportions of students in the following categories:

- Underachievers due to inadequate prior knowledge was relatively small at about 8% only.
- Underachievers due to over-assumed capabilities but with reasonable prior knowledge, as shown by students from the medium- and high-ability groups together, was large at about 63%.
- Students who achieved test scores as they expected was about 12 %.
- Students who achieved test scores beyond what they expected (or those with lower self-efficacy expectancies) was about 17%.

This study, thus, identified students who needed remedial counselling so that their achievement could be improved in the future by an appropriate teacher intervention program. If unchecked, the tendency to over-assume one's own judgemental capabilities might be replicated by students not only in mathematics but also across the curriculum. A list of students who over-assumed was communicated to teachers in different schools for a review.

Thus, my research study informs teaching practice about procedures to identify underachievers and provide remedial counselling for improving students' learning and achievement in the future.

Research objectives 4(c) and 4(d) of Subsection 1.5.2 involved investigating the predictive validity of the MSES by an analysis of correlations between efficacy-expectancies and achievement, and gender differences in efficacy-expectancies. Students' efficacy-expectancies were moderately correlated with their achievement (postscores), stronger for individual classes than the overall sample, and slightly lower for girls (0.39) than for boys (0.43), and thus gender differences in correlations were negligible.

For individual schools, however, there were significant gender differences in correlations between efficacy-expectancies and achievement because the correlations were higher for boys in Schools 2 and 3 and for girls in School 1 only. For other schools, gender differences were nonsignificant. Comparatively in science domain, Velayutham (2012) reported that, "...for both girls and boys, the most influential motivational belief on students' self-regulation is self-efficacy followed by learning goal orientation. However, although for boys the influence of task value was significant, for girls, this construct appeared to have a limited impact on their self-regulation in science learning" (pp. 139-140).

In my study, the effect sizes (Cohen's *d*) for gender differences in efficacy-expectancies suggest that boys of Schools 1, 2 and 6 had somewhat higher efficacy-expectancies than girls, whereas girls of Schools 3, 4 and 5 displayed efficacy-expectancies that were marginally higher than those of boys, but gender differences were nonsignificant in all schools.

These findings were consistent with those of Kruger (1999), who examined Year 9 science achievement for the multiple-choice and the constructed-response (short-answer) items and revealed the absence of important gender differences in achievement. However, his study, supported by qualitative analyses, reported more item-specific self-efficacy for boys than girls for both multiple-choice and short-answer test formats.

The MSES validated in my study can be used by teachers to predict students' achievement in their classes and inform teachers about the relative importance of task types, so that students' learning could be improved in different classes in future. By using a similar approach, teachers might improve their students' learning, with remedial counselling provided to students about exercising appropriate caution in self-reporting self-efficacy judgements.

8.2.5 Variance Component Analysis, Mediation Model and Two-Level SEM

My research objectives 5(a) and 5(b) (Subsection 1.5.2) involved the use of hierarchical data for estimation of direct and mediated effects of prior knowledge (PK) and

classroom learning environment (CLE) on achievement (ACH), with self-efficacy (SEI) as the mediator from a mediation model (MacKinnon et al., 1997). I estimated the joint influence of PK and CLE on ACH using a Two-level structural equation model (SEM).

As the data were hierarchical, I conducted a variance component analysis of the dependent variables, SEI and ACH, by following Goldstein (2011), and estimated that the intraclass variation was much higher at 64.1% for ACH and above 23% for SEI for the overall sample of 464 students. The corresponding estimates were much higher for boys than girls. Also, the intraclass variation in achievement postscores was much higher at 70.5% for the high-ability group, 59.2% for the medium-ability group and the least at 45.3% for the low-ability group, which revealed that there were wide BETWEEN Class differences in achievement within the high-ability group.

The ML estimates of variances in grand means indicated an important aspect of underachievement by boys and girls. Students of the higher-ability group displayed higher average self-efficacy expectancy (grand mean of SEI), which implies that many students of the high-ability group over-assumed their capabilities for successfully solving mathematics tasks on the posttest, which mirrors a similar finding reported in Subsection 6.3.2. It can therefore be inferred that, by an appropriate teacher intervention, students can be offered remedial counselling to verify students' (mis) understanding and improve their learning.

To accomplish research objective 5(a), the direct and mediated effects of prior knowledge and classroom learning environment on achievement were derived from the mediation model. Classroom learning environment as perceived by high-ability students had a greater direct effect (β=0.314) on achievement than perceived by low- and medium-ability groups together (β=0.065), which was not statistically significant. For the overall sample, the mediated effect of classroom learning environment was less on achievement at 0.121SD units than the corresponding direct effect (β = 0.219SD).

In contrast, the direct effect of prior knowledge of all types of tasks combined was statistically significant and higher for the low- and medium-ability groups combined (*b* =0.503 or β=0.353) than the high-ability group (b=0.422 or β=0.441). For the overall

sample, the mediated effect of prior knowledge was found to be smaller (b =0.114) than the corresponding direct effect (b =0.376).

A comparison of the direct effect of prior knowledge of task types P1 (PK1) with those of type P2 (PK2) revealed that the former type had a greater impact on achievement than the latter. Also, results suggested that the direct effects of both CLE and PK on achievement were higher for the higher-ability group than those for the overall sample.

To address research objective 5(b) of Subsection 1.5.2, the research model proposed earlier in Figure 1.1 was specified in three variants, SEM1 - SEM3:

- SEM1 posits the joint influence of the construct of classroom learning environment (CLE) and students' procedural knowledge for solving tasks of type P1 (PK1) on students' achievement (ACH) through the mediator of self-efficacy expectancy (SEI)
- SEM2 models the joint influence of CLE and students' procedural knowledge of solving mathematics tasks of type P2 (PK2) on ACH through the mediator, SEI
- SEM3 hypothesises the joint influence of CLE and students' prior knowledge of solving mathematics tasks all types (PK) on ACH through the mediator, SEI.

The results of hypotheses testing of joint influences were compared between variants and evaluated using important statistics such as the coefficient of determination (R^2), residual variance, Log L and the model fit information including the Comparative Fit Index (CFI) and the Akaike Information Criterion (AIC). Based on detailed evaluation, model SEM 3 was found to be the best representation of the joint influence of (CLE, PK) on students' achievement.

A major finding was that, for a mathematics topic of measurement that teachers taught to high-ability Year 10 students, the joint influence of classroom learning environment (CLE) and prior knowledge (PK) on achievement was favourable and significant, with the effect of PK being $\beta = 0.308$SD and that of CLE being $\beta = 0.148$ SD. For the whole sample, however, when estimated jointly from the SEM, only the effect of PK was statistically significant, but that of CLE was not.

Comparing the effects amongst the three variants, the joint influences of prior knowledge and classroom learning environment were greater for SEM 3 than the other two

variants, when prior knowledge of all types combined (PK) was used together with CLE, with their effects mediated significantly by self-efficacy on achievement. But, in any other SEM variant, the influence of CLE was not statistically significant although its mediated effect was significant.

A comparison of estimates from the two sub-samples of high-ability students (N=303) and the whole sample (N=464) indicated that high-ability students perceived their classroom learning environment as important and significant for improving achievement to the extent of 0.148 SD rather than all students in the whole sample.

Students' prior knowledge required for learning influenced achievement favourably and statistically significantly to the extent of b = 0.250 for the whole sample, and to the extent of b = 0.300 for the high-ability group. This was found to be relatively less than the corresponding direct effect of b = 0.376 for the overall sample and b = 0.422 for the high-ability group, probably due to covariance between prior knowledge and classroom learning environment.

The influence of CLE, obtained jointly with prior knowledge and as mediated by self-efficacy expectancy, was also statistically significant and greater in magnitude at β=0.247 SD for the high-ability group than for the overall sample (β=0.192 SD).

Thus, to sum up, classroom learning environment had a nonsignificant effect on achievement when students did not have adequate prior knowledge required for learning, which is plausible because prior knowledge was found to covary with classroom learning environment and had a bivariate correlation of 0.35 with achievement.

In general, the estimates of joint influence tended to be smaller in magnitude than their corresponding estimates of direct and mediated effects from the mediation model, which can be attributed to possible covariances between the explanatory variables included in the SEM. The estimates of joint influence thus offer a closer approximation to reality for gauging the joint impact of students' prior knowledge and classroom learning environment on achievement.

8.3 Significance and Contributions of Study

This study bridged a research gap within the field of learning environment by examining the joint influence of prior knowledge and classroom learning environment on students' self-efficacy beliefs and achievement in mathematics learning. My study's theoretical contribution lies in identifying prior knowledge as an important determinant that covaries with classroom learning environment and thereby adding to the literature on students' achievement in mathematics learning. It benefits both the fields of learning environment and educational psychology by setting a precedent for future studies.

The first methodological contribution of this study is that the COLES was modified to form the MCOLES, which is suitable for mathematics classes. To establish the validity of the newly-developed instrument, a comprehensive construct validity framework was employed involving face, convergent, discriminant, concurrent and predictive validities. The validated MCOLES has 50 items only and was used for measuring students' perceptions of the construct of classroom learning environment. This method could be used by future researchers who wish to develop and validate new questionnaires.

The second methodological contribution was in the extraction of factors from MCOLES responses, as well as an application of second-order CFA to obtain a parsimonious representation of the construct by a single factor. This is analytically amenable for estimating the direct and mediated effects of classroom learning environment using a mediation model, and jointly with prior knowledge on achievement from the SEM.

Third, this study also developed a classroom tool for calibrating students' efficacy judgements into efficacy-expectancies, measured by students' expected scores in the achievement posttest. It thus demonstrated how teachers could use a similar procedure for identifying underachievers and those students who display lower self-efficacy in successfully achieving mathematics tasks of any topic taught. Thus, this tool can be used by teachers when considering appropriate classroom interventions for improving the future achievement of boys and girls, separately.

Fourth, major contributions emerged from estimation of effect sizes and gender differences in students' prior knowledge and achievement, as well as differences in correlations between students' achievement and i) prior knowledge, ii) classroom learning environment and iii) self-efficacy expectancies, which add to the literature of mathematics education research in secondary schools. The results of direct and mediated effects of students' prior knowledge and classroom learning environment and estimates of their joint influence from Two-level SEM of achievement could be used in future studies to derive information related to the mediational role of self-efficacy.

Finally, the fifth contribution of this study was that researchers and teachers were provided with some convenient tools to help with organising students in mixed-ability classroom into low-, medium- and high ability groups based on their prior knowledge, and with adopting teaching practices that offer equally-challenging learning opportunities for students of all ability levels and genders. Teachers could use these tools for refocussing their pedagogical approaches and evaluating instructional strategies that have the potential to increase the achievement of both boys and girls.

8.4 Limitations

The first limitation was about the sample and drawing inferences based on a purposive sample from five Indian schools and one Australian school, which disproportionately represents the underlying student populations in these two countries at the Year 10 level. No comparisons were made between India and Australia in classroom teaching practices or students' learning based on sample findings. However, given the objectives of the study, generalised inferences were made using the sample-based findings about the effects of cognitive and noncognitive factors on achievement and other analytical aspects such as:

i) development, validation and factor analysis of the new instrument, MCOLES

ii) hypothesis testing of gender differences in students' prior knowledge and achievement

iii) hypothesis testing for changes in learning

iv) direct and mediated effects of self-efficacy beliefs

v) path coefficients from two-level SEM.

Second, my sample finally used 464 complete records, which included only 51 observations (student records) from the low- and 160 from the medium- and 303 high-ability groups from a total 15 classes; these sample sizes were inadequate for investigating Two-level SEM analysis for the low- and medium-ability groups separately. Thus, the analysis was confined to the whole sample (N=464) and the high-ability group (N=303) only.

Similarly, the joint influence could not be assessed separately for boys and girls, or low- and medium-ability groups, because of sample size limitations (i.e. the number of 15 classes in the sample of 464 students restricted the researcher to including only one explanatory variable at a time in running Level 2 regressions). The explanatory power of the SEM could be improved with a sample having a larger number of classes and more explanatory variables as suggested below.

The sample was somewhat restrictive when obtaining maximum likelihood estimates of parameters involved in the second level of SEM because of limited degrees of freedom. As a result, at Level 2, only one explanatory variable could be used at a time for explaining the variance in students' achievement BETWEEN Classes, which therefore resulted in low explanatory power for the model.

Third, there was measurement error in the class size variable in the sample. Class size, which was measured by the number of students who provided all necessary information required for the study, ranged from 16 students in Class 5 to 41 students in Classes 4 and 6. These numbers were less than the actual number of students attending classes because some of those who were enrolled did not complete all necessary surveys and tests in this study. Many such students who were excluded from the analysis attended

classes and probably influenced the teaching time otherwise available for each student. Also, their presence in the class might have impacted their classroom learning environment and, to an extent, their learning outcomes, which could not be segregated in this analysis, and thus this remains to be a limitation.

The fourth limitation concerns the explanatory power of the Two-level SEM estimated. The variance explained in the criterion variable (ACH) was approximately up to 40% in the WITHIN Class component and up to 12 % only in the BETWEEN Classes component. There was still a substantial degree of unexplained variance in random intercepts in ACH and SEI in the model.

Finally, the fifth limitation relates to calibrating students' judgements using the MSES responses, with an assumption of scale continuity between any two response opportunities available to students. When students expected a score between any two consecutive options, only the nearest option was available, which could be higher or lower than what student would have preferred. As a result, it could penalise students incorrectly by classifying them as underachievers. Bandura (2006) recommended that scales be designed with response opportunities as close as possible to minimise such error. This study allowed for ± 5% error.

8.5 Suggestions for Future Research

To address the issue of diversity in mixed-ability classrooms in this research, teachers taught students after dividing them into three ability-groups based on their prior knowledge scores, but the number of groups into which they divided the whole class was rather arbitrary. A larger number of groups requires more of the teachers' time in managing them. Thus, future research might address how to divide students and allocate work to engage them effectively. Alternative criteria for grouping could include the coefficient of variation in prescores that reflects the extent of teaching challenge involved, as well as

teachers' time spent for lesson planning by selection of mathematics tasks from the curriculum to cater for the cognitive needs of students. Teachers might seek an optimal solution that minimises teaching challenge and time and effort invested for teaching ability groups while maximising students' achievement.

To enhance the explanatory power of the Two-level SEM, future research should involve a larger number of classes and other explanatory variables such as student cooperation, peer tutoring, clubs, parental qualification and its influence on homework, and alternative measures of class size. For a higher-level modelling such as Level 3, students' socio-economic status might be incorporated to explain variance as BETWEEN schools in students' achievement and self-efficacy beliefs, covering the participation of an appropriate number of schools in future studies.

One might examine variance in students' achievement and self-efficacy beliefs under the assumption of random slopes and/or random intercepts models.

The predictive validity of the MCOLES could be further explored by examining associations between the construct of classroom learning environment and other learning outcomes, such as students' attitudes towards mathematics and test anxiety. Also, it would be possible to use the MCOLES to evaluate teaching-related innovations involving new methods of grouping students, peer tutoring in terms of their impact on the classroom environment (Fraser, 2012). Finally, the MCOLES could be gainfully employed for teacher action research aimed at improving mathematics classrooms by considering students' preferences about different dimensions of classroom learning environment (Aldridge et al., 2012).

8.6 Final Comment

This study has achieved all its research objectives, provided important theoretical, methodological and practical contributions and encouraged possible future research

directions. The theoretical contributions add to the literature on mathematics education whilst the methodological contributions present viable alternatives options for future researchers. The practical implications alert educators to plan and put into practice effective pedagogical strategies aimed at increasing students' self-efficacy beliefs and achievement. Because the focus of this research was prior knowledge required for mathematics learning, the findings probably could help educators to understand and improve students' motivational self-efficacy beliefs by using the classroom tools developed and demonstrated.

There has been a shift of emphasis from teachers' instructional techniques to developing students' self-efficacy beliefs in learning. The teacher's role today is to increase students' motivation and develop necessary skills or strategies that help to make them responsible learners, with a focus on restructuring students' learning environment so that they can take ownership of their own learning.

The findings of the study highlight the role of self-efficacy expectancies in mediating the effects of students' prior knowledge and classroom learning environment on achievement. These valuable insights could be shared to fill future mathematics classrooms with motivated learners and with enhanced self-efficacy beliefs in achieving.

REFERENCES

ACARA. (2009). *Shape of the Australian curriculum: Mathematics*. Retrieved from http://www.acara.edu.au/verve/_resources/Australian_Curriculum_-_Maths.pdf.

ACARA. (2012). *NAPLAN summary report*. Retrieved from www.acara.edu.au

ACARA. (2015). *Australian curriculum/mathematics*. Retrieved from http://www.australiancurriculum.edu.au/

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csáki (Eds.), *2nd International symposium on information theory* (pp. 267–281). Budapest: Akadémiai Kiadó.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.

Akaike, H. (1985). Prediction and entropy. In A. C. Atkinson & S. E. Fienberg (Eds.), *A celebration of statistics* (pp. 1-24). Dordrecht, the Netherlands: Springer.

Albert, L. R. (2012). *Rhetorical ways of thinking: Vygotskian theory and mathematical learning.* Dordrecht, the Netherlands: Springer.

Aldridge, J. M., & Fraser, B.J. (2008). *Outcomes-focused learning environments: Determinants and effects*. Rotterdam, The Netherlands: Sense.

Aldridge, J. M., Fraser, B. J., Bell, L., & Dorman, J. (2012). Using a new learning environment questionnaire for reflection in teacher action research. *Journal of Science Teacher Education, 23*(3), 259-290. http://dx.doi.org/10.1007/s10972-012-9268-1

Aldridge, J. M., Fraser, B. J., Taylor, P. C., & Chen, C. C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22*(1), 37-55.

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315-343. http://dx.doi.org/10.3102/00346543061003315

Alreck, P. L., & Settle, R. B. (1995). *The survey research handbook: Guidelines and strategies for conducting a survey* (2nd ed.). Chicago: Irwin Professional Publishing.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (5th ed.). New York: Wroth.

Askew, M. (2015). Diversity, inclusion and equity in mathematics classrooms: From individual problems to collective possibility. In A. Bishop, H. Tan, & T. N. Barkatsas (Eds.), *Diversity in mathematics education: Towards inclusive practices* (pp. 129-145). Cham, Switzerland: Springer International Publishing.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.

Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. London: Kluwer Academic Publishers.

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

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

Bandura, A. (1980). Gauging the relationship between self-efficacy judgment and action. *Cognitive Therapy and Research, 4*(2), 263-268. http://dx.doi.org/10.1007/bf01173659

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Bandura, A. (1986a). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986b). The explanatory and predictive scope of self-efficacy theory. *Journal of Clinical & Social Psychology*, *4*, 359-373.

Bandura, A. (1997). *Self-efficacy*. New York: W. H. Freeman

Bandura, A. (2006). Guide for creating self-efficacy scales. In F. Pajares, & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-338). Charlotte, NC: Information Age Publications.

Bandura, A., Reese, L., & Adams, N. E. (1982). Microanalysis of action and fear arousal as a function of differential levels of perceived self-efficacy. *Journal of Personality and Social Psychology, 43*(1), 5-21. http://dx.doi.org/10.1037/0022-3514.43.1.5

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16 (Series B)*, 296-298.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting & Clinical Psychology, 42*(2), 155-162.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246. http://dx.doi.org/10.1037/0033-2909.107.2.238

Bloom, B.S. (1976). *Human characteristics and school learning*, New York: McGraw-Hill.

Brooks, J. G. (1990). Teachers and students: Constructivists forging new connections. *Educational Leadership, 47*(5), 68-71.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Publications.

Broza, O., & Kolikant, Y. (2015). Contingent teaching to low-achieving students in mathematics: Challenges and potential for scaffolding meaningful learning. *The International Journal of Mathematics Education, 47*(7), 1093-1105. http://dx.doi.org/10.1007/s11858-015-0724-1

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276. doi:10.1207/s15327906mbr0102_10

Chhapia, H. (2012, January 15). Indian students rank 2nd last in global test. *The Times of India.* Retrieved from http://timesofindia.indiatimes.com/home/education/news/Indian-students-rank-2nd-last-in-global-test/articleshow/11492508.cms

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology, 47*(5), 291-314. http://dx.doi.org/10.1016/j.jsp.2009.04.002

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Collins, L. (1982, March). *Self-efficacy and ability in achievement behaviour.* Paper presented at the Annual meeting of the American Educational Research Association, New York.

Creswell, J. W., & Plano Clark, V. L. (2007). The nature of mixed methods research. In J. W. Creswell, & Plano Clark, V. L (Eds.), *Designing and conducting mixed methods research* (pp. 1-18). Thousand Oaks, CA: Sage.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Crotty, M. (1998). *Foundations of social research: Meaning and perspective in the research process.* Sydney, NSW: Allen & Unwin

Davis, E. W., & Yates, B. T. (1982). Self-efficacy expectancies versus outcome expectancies as determinants of performance deficits and depressive affect. *Cognitive Therapy and Research, 6*, 23-35.

De Ayala, R. J. (2009). *The theory and practice of item response theory*, New York: Guilford Press.

Dochy, F. J. R. C. (1992). *Assessment of prior knowledge as a determinant for future learning: The use of prior knowledge state tests and knowledge profiles*. Utrecht: Lemma

Dochy, F. J. R. C. (1994). Prior knowledge and learning. In T. S. Husen, & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed., pp. 4698-4702). London/New York: Pergamon.

Dochy, F. J. R. C. (1996). Assessment of domain-specific and domain-transcending prior knowledge: Entry assessment and the use of profile analysis. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 227-264). Boston: Kluwer.

Dochy, F.J.R.C., & Alexander, P.A. (1995). Mapping prior knowledge: A framework for discussion among researchers. *European Journal of Psychology of Education*, *10*, 225-242.

Dochy, F.J.R.C., De Ridjt, C., & Dyck, W. (2002). Cognitive prerequisites and learning: How far have we progressed since Bloom? Implications for educational practice and teaching. *Active Learning in Higher Education*, *3*, 265–284.

Dorman, J. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research*, *4*(3), 243-257. http://dx.doi.org/10.1023/a:1014490922622

Dorman, J., & Adams, J. (2004). Associations between students' perceptions of classroom environment and academic efficacy in Australian and British secondary schools. *Westminster Studies in Education*, *27*(1), 69-85. http://dx.doi.org/10.1080/0140672040270106

Dowling, D. M. (1978). *The development of a mathematics confidence scale and its application in the study of confidence in women colleges.* Retrieved from https://search.proquest.com.dbgw.lis.curtin.edu.au/docview/302901047?accountid=10382

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103-127. http://dx.doi.org/10.1037/a0018053

Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.

Ernest, P. (1991). *The philosophy of mathematics education*. London, UK: The Falmer Press.

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. Albany, NY: SUNY Press.

Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., & Hammond, K. A. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *Journal of Educational Psychology, 102*(3), 729-740.

Fennema, E., & Leder, G. C. (1990). *Mathematics and gender.* New York: Teachers College Press

Fennema, E., & Sherman, J. (1976). Fennema-Sherman mathematics attitudes scales. *JSAS Catalog of Selected Documents in Psychology, 6*(1), 31.

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal, 14*(1), 51-71.

Fisher, R. A. (1924). The distribution of the partial correlation coefficient. *Metron, 3*(3-4), 329-332.

Forgasz, H. (2010a). Streaming for mathematics in Victorian secondary schools. *Australian Mathematics Teacher, 66*(1), 31-40.

Forgasz, H. (2010b). Streaming for mathematics in years 7–10 in Victoria: An issue of equity? *Mathematics Education Research Journal, 22*(1), 57-90.

Forgasz, H., & Leder, G. (2015). Surveying the public: Revisiting mathematics and English stereotypes. In A. Bishop, H. Tan & T. N. Barkatsas (Eds.), *Diversity in mathematics education: Towards inclusive practices* (pp.103-117). Cham, Switzerland: Springer International Publishing

Forgasz, H. J., & Leder, G. C. (2017). Persistent gender inequities in mathematics achievement and expectations in Australia, Canada and the UK. *Mathematics Education Research Journal, 29*(3), 261-282. doi:http://dx.doi.org/10.1007/s13394-017-0190-x

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* New York: McGraw-Hill.

Fraser, B. J. (1990). *Individualized Classroom Environment Questionnaire*. Melbourne, Victoria: Australia: Australian Council for Educational Research.

Fraser, B. (1998). Classroom environment instruments: Development, validity and applications. *Learning Environments Research, 1*(1), 7-34.

Fraser, B. (2002). Learning environments research: Yesterday, today and tomorrow. In S. C. Goh, & Khine, M.S (Eds.), *Studies in educational environments: An international perspective* (pp. 1-25). Singapore: World Scientific.

Fraser, B. (2007). Classroom learning environments. In S. K. Abell, & N.G. Lederman (Eds.), *Handbook of research on science education* (pp. 103-124). London, UK: Lawrence Erlbaum Associates.

Fraser, B. (2012). Classroom learning environments: Retrospect, context and prospect. In B. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191-1239). Dordrecht, The Netherlands: Springer.

Fraser, B. J. (2014). Classroom learning environments: Historical and contemporary perspectives. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. II, pp. 104-119). New York: Routledge.

Fraser, B., & Tobin, K. (1991). Combining qualitative and quantitative methods in classroom environment research. In B. Fraser, & H. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 271-292). Oxford, UK: Pergamon Press.

Germann, P. (1994). Testing a model of science process skills acquisition: An interaction with parents' education, preferred language, gender, science attitude, cognitive development, academic ability, and biology knowledge. *Journal of Research in Science Teaching, 31*, 749-783.

Giele, J. (1988). Gender and sex roles. In N. Smelser (Ed.), *Handbook of sociology*. Newbury Park, CA: Sage.

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). New York: John Wiley.

Government of India. (2005). *Statistical abstract, India*. New Delhi, India: Central Statistical Organisation, Ministry of Statistics & Programme Implementation.

Government of India. (2015). *Statistical abstract, India*. New Delhi, India: Central Statistical Organisation, Ministry of Statistics & Programme Implementation.

Guba, E. G. (Ed.). (1990). *The paradigm dialog*. Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin, & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.

Gupta, A., & Fisher, D. (2012). Technology-supported learning environments in science classrooms in India. *Learning Environments Research, 15*(2), 195-216.

Gurría, A. (2016). *PISA 2015 results in focus*. Paris: Organisation for Economic Cooperation and Development. Retrieved from https://search.proquest.com/docview/1853270009?accountid=10382

Hackett, G. (1985). Role of mathematics self-efficacy in the choice of math-related majors of college women and men: A path analysis. *Journal of Counselling Psychology, 32*(1), 47-56.

Hackett, G., & Betz, N. E. (1981). A self-efficacy approach to the career development of women. *Journal of Vocational Behavior, 18*(3), 326-339.

Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/ mathematics performance correspondence. *Journal for Research in Mathematics Education, 20,* 261-273.

Hailikari, T., Katajavuori, N., & Lindblom-Ylanne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American Journal of Pharmaceutical Education, 72*(5), 113.

Hailikari, T., Nevgi, A., & Komulainen, E. (2007). Academic self-beliefs and prior knowledge as predictors of student achievement in Mathematics: A structural model. *Educational Psychology, 28*(1), 59-71.

Hailikari, T., Nevgi, A., & Lindblom-Ylänne, S. (2007). Exploring alternate ways of assessing prior knowledge, its components and their relation to students achievement *Studies in Educational Evaluation, 33*(3-4), 320-337. http://dx.doi.org/10.1016/j.stueduc.2007.07.007

Hallinan, M. T., & Sorensen, A. B. (1987). Ability grouping and sex differences in mathematics achievement. *Sociology of Education, 60*(2), 63-72.

Harlow, S., Cummings, R., & Aberasturi, S. M. (2006). Karl Popper and Jean Piaget: A rationale for constructivism. *The Educational Forum, 71*(1), 41-48.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York : Routledge

Henderson, D., Fisher, D. L., & Fraser, B. J. (1995). *Associations between learning environments and student outcomes in biology*. Paper presented at the Annual Meetings of the American Educational Research Association, San Francisco, CA.

Hill, C., Corbett, C., & St. Rose, A. (Eds.). (2010). *Why so few? Women in science, technology, engineering, and mathematics* Washington, DC: American Association of University Women (AAUW). Retrieved from http:// www. aauw. org/ learn/ research/ upload/ whysofew. pdf

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14306381

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76– 99). Thousand Oaks, CA: Sage.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453.

Hu, L., & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.

Hubbard, R., & Allen, S. J. (1987). An empirical comparison of alternative methods for principal component extraction. *Journal of Business Research, 15*, 173-190.

Jaeger, R. M. (1997). Survey research methods in education. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 449-476). Washington, DC: American Educational Research Association.

Kaiser, H. (1970). A second-generation little jiffy. *Psychometrika, 35*, 401-415.

Kaiser, H. (1974). An index of factorial simplicity. Psychometrika, 39, 31-36.

Kendall M. G., & Stuart, A. (1973). *The advanced theory of statistics* (3rd ed.). New York: Hafner.

Kenny, D.A. (2015). Introduction to Fit Indices. Retrieved from http://davidakenny.net/cm/fit.htm

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486-507

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*, 333-351.

Koul, R. B. (2003). *Teacher-student interactions and science classroom learning environments in India.* Unpublished Doctoral thesis, Curtin University of Technology, Perth, WA.

Koul, R. (2009). *Indian classroom learning environments: Culture and education*. Saarbrucken, Germany: VDM Verlag.

Krueger, B. (1999). *Gender and test item-response formats.* Doctor of Science Education (Unpublished Thesis), Curtin University, Perth.

Kuhn, T. K. (1962). *The structure of scientific revolutions* Boston, MA: Chicago University Press.

Law, A. (2014). *The effect of tracking on student mathematical achievement growth at a select school district in suburban Chicago, Illinois.* (D.Ed.), Aurora University, Ann Arbor. Retrieved from https://search.proquest.com/docview/1566943262?accountid=10382

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data.* New York: John Wiley & Sons.

Locke, E. A. (1987). Social foundations of thought and action: A social-cognitive view. *Academy of Management Review, 12*(1), 169-171.

Lowry, R. (2017). Concepts & applications of inferential statistics. Retrieved from http://vassarstats.net/index.html

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

MacKinnon, D. P. (2007). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.

MacKinnon, D. P., & Dwyer J.H. (1993). Estimation of mediation effects in prevention studies. *Evaluation Review, 17*, 144-158.

MacKinnon, D. P., Dwyer J.H., & Warsi, G. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research, 30*, 41-62.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*(1), 593-614.

Maddux, J. E. (Ed.). (1995). *Self-efficacy, adaptation, and adjustment: Theory, research, and application*. New York: Plenum Press.

Maddux, J. E., Norton, L. W., & Stoltenberg, C. D. (1986). Self-efficacy expectancy, outcome expectancy, and outcome value: Relative effects on behavioral intentions. *Journal of Personality & Social Psychology, 51*(4), 783-789.

Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology*, *19*(5), 469-479. http://dx.doi.org/https://doi.org/10.1016/0022-1031(83)90023-9

Marsh, H. W., Craven, R. G., & McInerney, D. (Eds.). (2005). *New frontiers in SELF research.* Charlotte, NC: Information Age Press.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Midgley, C., & Edelin, K. C. (1998). Middle school reform and early adolescent well-being: The good news and the bad. *Educational Psychologist*, *33*, 195-206.

Midgley, C., & Urdan, T. (1995). Predictors of middle school students' use of self-handicapping strategies. *Journal of Early Adolescence*, *15*, 389-411.

Moll, L. C., Estrada, E., Diaz, E., & Lopes, L. M. (1980). The organization of bilingual lessons: Implications for schooling. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, *2*(3), 53-58.

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

Moos, R. H. (1979). *Evaluating educational environments: Procedures, measures, findings and policy implications.* San Francisco, CA: Jossey Bass.

Munby, H. (1997). Issues of validity in science attitude measurement. *Journal of Research in Science Teaching*, 34, 337-341.

Murphy, P., & Gipps, C. V. (1996). *Equity in the classroom: Towards effective pedagogy for girls and boys.* Washington, DC: Falmer Press.

Muthen, L. K., & Muthen, B. (2008). M*plus user's guide*. Los Angeles, CA: Author.

NCERT. (2012). *Class X mathematics.* Retrieved from www.ncert.nic.in

NCERT. (2015). *National curriculum framework - 2005*. Retrieved from http://epathshala.nic.in/wp-content/doc/NCF/Pdf/nf2005.pdf

Niemann, D., Martens, K., & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education, 52*(2), 175-183. http://dx.doi.org/10.1111/ejed.12220

Niemi, H., Nevgi, A., & Virtanen, P. I. (2003). Towards self-regulation in web-based learning. *Journal of Educational Media, 28*(1), 49-71. doi:10.1080/1358165032000156437

Norwich, B. (1986). Assessing perceived self-efficacy in relation to mathematics tasks: A study of the reliability and validity of assessment. *British Journal of Educational Psychology, 56*(2), 180-189.

Oakes, J. (1992). Can tracking research inform practice? Technical, normative, and political considerations. *Educational Researcher, 21*(4), 12-21.

Oakes, J. (2005). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

OECD. (2016a). *Are low performers missing learning opportunities?* Paris, FR: OECD Publishing. Retrieved from /content/working paper/5jlz6r8lvwwc-en http://dx.doi.org/10.1787/5jlz6r8lvwwc-en.

OECD. (2016b). *Who are the low-performing students?* Paris, FR: OECD Publishing.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443. http://dx.doi.org/10.1006/ceps.1995.1029

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology, 42*(2), 190-198.

Pajares, F., & Urdan, T. C. (Eds.). (2006). *Self-efficacy beliefs of adolescents.* Greenwich, CT: Information Age Publishers.

Pallant, J. (2013). *SPSS survival manual*. Sydney, NSW, Australia: Allen & Unwin

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*(3), 293-316.

Parker, L. H., Rennie, L. J., & Fraser, B. J. (Eds.). (1995). *Gender, science and mathematics: Shortening the shadow.* Dordrecht: Kluwer Academic Publishers.

Patel, S. (2015). *The research paradigm-methodology, epistemology and ontology – explained in simple language*. Retrieved from http://salmapatel.co.uk/academia/the-research-paradigm-methodology-epistemology-and-ontology-explained-in-simple-language

Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *Journal of Educational Psychology, 85*(1), 41-48.

Raudenbush, S. W., & Sampson, R. (1999). Assessing direct and indirect effects in multilevel designs with latent variables. *Sociological Methods Research, 28*, 123-153.

Rawnsley, D. G. (1997). *Associations between classroom learning environments, teacher interpersonal behaviours and student outcomes in secondary mathematics classrooms* (Unpublished Ph.D). Curtin University of Technology, Perth, WA.

Rizvi, F., & Gorur, R. (2011). Collaboration in higher education: New directions for India and Australia: An issues paper for the Australia India Education Council. Retrieved from http://www.australiaindiaeducation.com

Rizvi, F., Gorur, R., & Reyes, C. (2015). India-Australia institutional collaborations in higher education: Potential, problems, promises. Retrieved from http://www.australiaindiaeducation.com/news-2.html

Roeser, R. W., Midgley, C., & Urdan, T. (1996). Perception of the school psychological environment and early adolescents' self-appraisals and academic engagement: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*, 408-422.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29*(4), 304-321. http://dx.doi.org/10.1177/0734282911406653

Schneider, W., Korkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology, 81*, 306-312.

Schunk, D. H. (1995). Self-efficacy, education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and* adjustment: *Theory, research, and application* (pp.281-303). New York: Plenum Press.

Schunk, D. H. (2012). *Motivation and self-regulated learning: Theory, research, and applications*. Hoboken, NY: Taylor and Francis.

Schunk, D. H., & Miller, S. D. (2002). Self-efficacy and adolescents' motivation. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents* (pp. 29-52). Greenwich, CT: Information Age Publishers.

Schunk, D. H., & Pajares, F. (2004). Self-efficacy in education revisited: Empirical and applied evidence In D. McInerney & S. Van Etten (Eds.), *Research on sociocultural influences on motivation and learning* (Vol. 4, pp. 115-138). Greenwich, CT: Information Age Publishers.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461-464. doi:10.1214/aos/1176344136

Schwarzer, R. (2014). Self-efficacy and human functioning. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 1-39). London: Routledge, Taylor and Francis.

Siegel, R. G., Galassi, J. P., & Ware, W. B. (1985). A comparison of two models for predicting mathematics performance: Social learning versus math aptitude-anxiety. *Journal of Counseling Psychology, 32*(4), 531-538.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods,* New York: Routledge.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.

Sreekanth, Y., Tewari, A. D., Srivastava, S., & Bhushan, S. (2015). *What students know and can do: A summary of national achievement survey Class X, 2015.* New Delhi, India: National Council of Educational Research and Training, Educational Survey Division.

Stacey, K., & Turner, R. (Eds.). (2015). *Assessing mathematical literacy: The PISA experience.* Cham, Switzerland: Springer International Publishing.

State Government of Victoria. (2012). *Principles of learning and teaching*, Melbourne, VIC. Retrieved from http://www.education.vic.gov.au/school/teachers/support/Pages/unpacked12.aspx#anchor_four

State Government of Victoria. (2017). *High impact teaching strategies*. Retrieved from www.education.vic.gov.au/school/teachers/teachingresources/practice/improve/Pages/hits.aspx

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement. *Review of Educational Research, 86*(4), 849-899. http://dx.doi.org/doi:10.3102/0034654316675417

Steiger, J. H., & Lind, J. C. (1980). *Statistically-based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA. https://doi-org.dbgw.lis.curtin.edu.au/10.1080/10705511.2016.1217487

Sullivan, P. (2015). Maximising opportunities in mathematics for all students: Addressing within-school and within-class differences. In A. Bishop, H. Tan & T. N. Barkatsas (Eds.), *Diversity in mathematics education: Towards inclusive practices*. Cham, Switzerland: Springer International Publishing. http://dx.doi.org/:10.1007/978-3-319-05978-5_

Sullivan, P., & Mornane, A. (2014). Exploring teachers' use of, and students' reactions to, challenging mathematics tasks. *Mathematics Education Research Journal, 26*(2), 193–213.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Taylor, B. A., & Fraser, B. J. (2013). Relationships between learning environment and mathematics anxiety. *Learning Environments Research, 16* (2), 297-313. http://dx.doi.org/http://dx.doi.org/10.1007/s10984-013-9134-x

Taylor, P. C. (1994) Collaborating to reconstruct teaching: The influence of researcher beliefs. In K Tobin (Ed.), *The practice of constructivism in science education* (pp. 267– 298) Hillsdale, NJ: Lawrence Erlbaum.

Taylor, P. C. (1996). Myth making and myth breaking in the mathematics classroom. *Educational Studies in Mathematics, 31*, 151-173.

Taylor, P. C., Dawson, V., & Fraser, B. J. (1995) *A constructivist perspective on monitoring learning environments under transformation*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Taylor, P. C., & Fraser, B. J. (1991) *Development of an instrument for assessing constructivist learning environments*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293-302.

Taylor, P. C., Fraser, B. J., & White, L. R. (1994). *The revised CLES: A questionnaire for educators interested in the constructivist reform of school science and mathematics*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED Majors. *Sex Roles, 77*, 86-96. doi:10.1007/s11199-016-0694-y

Thompson, R.A., & Zamboanga, B.L. 2004. Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, 96, 778–784.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37-54.

Tobin K., & Fraser, B. J. (Eds.) (1998). Qualitative and quantitative landscapes of classroom learning environments. In B. J. Fraser and K. G. Tobin (Eds.), *The*

international handbook of science education (pp. 623-640). Dordrecht, The Netherlands: Kluwer Academic Publishers,

Trochim, W. M., & Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati, OH: Atomic Dog.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1– 10.

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology 44*, 331-349.

Valentine, J.C., Du Bois, D.L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*, 111–133.

Velayutham, S. D. (2012). *The influence of classroom environment on students' motivation and self-regulation* (Unpublished PhD). Curtin University of Technology, Perth, WA.

Velayutham, S., Aldridge, J., & Fraser, B (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation structural modelling analysis. *International Journal of Science and Mathematics Education, 10*(6), 1347-1368. doi:10.1007/s10763-012-9339-y

von Glasersfeld, E. (1985). Constructivism in education. In T. Husen & T. N. Postlethwaite (Eds.), *The International encyclopedia of education: Research and studies* (Vol. 1, pp. 162-163). Oxford, UK: Pergamon Press.

von Glasersfeld, E. (2000). Problems of constructivism. In L. P. Steffe & P. W. Thompson (Eds.), *Radical constructivism in action: Building on the pioneering work of Ernst Glasersfeld* (pp. 3-9). New York: Routledge.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walshaw, M. (2016). Lev Vygotsky. In E. de Freitas & M. Walshaw (Eds.), *Alternative theoretical frameworks for mathematics education research: Theory meets data* (pp. 11-37). Cham, Switzerland: Springer International Publishing.

Watkins, M. W. (2000). *Monte Carlo PCA for parallel analysis*. State College, PA: Ed & Psych Associates.

Weinert, F. E., & Helmke, A. (1998). The neglected role of individual differences in theoretical models of cognitive development. *Learning and Instruction, 8*(4), 309-323. http://dx.doi.org/http://dx.doi.org/10.1016/S0959-4752(97)00024-8

Wood, D., Bruner, J. C., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100.

Wood, T., & Turner-Vorbeck, T. (2001). Extending the conception of the mathematics teaching. In T. Wood, B. S. Nelson & J. Warfield (Eds.), *Beyond classical pedagogy* (pp. 1895-1208). Mahwah, NJ: Lawrence Erlbaum.

Wu, M., Tsung, H. P. T., & Jen, H. (2016). *Educational measurement for applied researchers: Theory into practice*. Singapore: Springer Nature.

Yang, Y. & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment, 29*(4), 377-392. http://dx.doi.org/10.1177/0734282911406668

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

Websites

http://www.assessmentforlearning.edu.au/professional_learning/learning_intentions/learning_examples_intentions.html#3

http://www.australiaindiaeducation.com

http://www.australiancurriculum.edu.au/Mathematics/Rationale

http://www.davidakenny.net/cm/fit.htm

http://www.epathshala.nic.in/wp-content/doc/NCF/Pdf/nf2005.pdf).

http://www.google.co.in

http://www.ncert.nic.in

http://www.sathyasai.org/saieducation/content.htm

http://www.scholar.google.com

http://www.socscistatistics.com/pvalues/tdistribution.aspx

http://www.vassarstats.net/index.html

http://www.victoriancurriculum.vcaa.vic.edu.au

Appendix 1

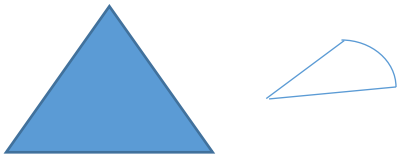
KNOWLEDGE FRAMEWORK

The operationalisation of different components of prior knowledge required for learning a unit of lessons on a topic *measurement* in Year 9 as per Australian mathematics curriculum: A sample *criterial test*

Components of knowledge	Operationalisation	Sample items

Declarative Knowledge

Type D1

Knowledge of facts meanings	Free recall; enumerating, Open question:	*Identify which of these two shapes is closer to an isosceles triangle and which one is closer to a sector.*
.	.	

Type D2

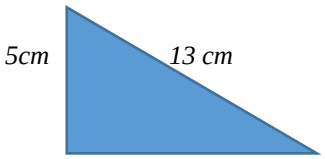
Knowledge of concepts, definitions		*The total surface area of a hemisphere is close to which of the following* to be recalled *formulae?*

a) $\pi r^2 h$
b) $2\pi r^2$
c) $3\pi r^2$
d) $2/3\pi r^3$

(contd.)

Appendix 1 (contd..2)

KNOWLEDGE FRAMEWORK

Components of knowledge	Operationalisation	Sample items
Procedural Knowledge		
Type P1		
Integration of knowledge between different mathematical concepts	Questions about interrelations	*Calculate the area and perimeter of the geometrical shape, given the measurements of its sides below.* 
Type P2		
Application of knowledge	Mathematical problem solving tasks	*A sports girl is practising a marathon race of 3.140 kms in a circular track of radius 100 m. How many rounds (at least) would she need to make in order to complete the race?* *a) Ten* *b) Five* *c) Twenty* *d) Fifty*

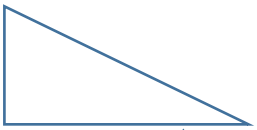
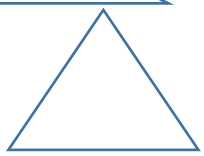
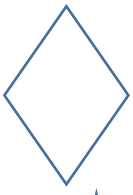
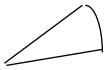
Notes: Columns 1 and 2 above were adapted from Hailikari, Nevgi and Lindblom-Ylänne (2007)

Appendix 2

PRIOR KNOWLEDGE TEST (PRETEST) FOR TOPIC OF MEASUREMENT

1. Match the geometrical shapes below by writing the corresponding number in the brackets

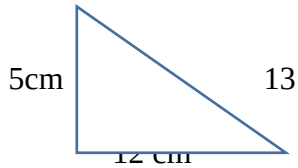
Marks 8 x ½ = 4

Geometrical shape	Name	
1 	A. Parallelogram	( )
2 	B. Rhombus	( )
3 	C. Right angled triangle	( )
4 	D. Equilateral triangle	( )
5 	E. Rectangle	( )
6 	F. Sector	( )
7 	G Isosceles triangle	( )
8 	H. Trapezium	( )

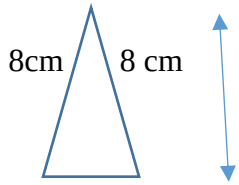
Appendix 2 (contd—2)

II. Calculate the area and perimeter of the following geometrical shapes using an appropriate formula.

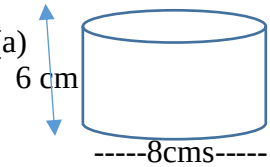
Marks

1 

Area Perimeter 2

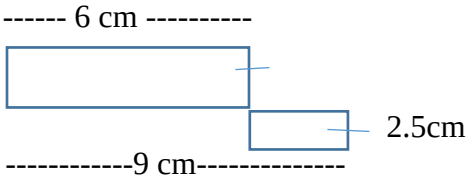
2 

Area Perimeter 2

3 (a) 

i) Curved surface area=**….** cm^2 ii) Top area= **……**cm^2 4

iii) Total surface area = cm^2 iv) Volume =……cm^3

(b) 

Area Perimeter 2

III. If your mathematics text book size is given by 16cm x 12cm x 8cm, where length= 16cm; width= 12 cms and height = 8cms, work out the following,

a) Total surface Area 1

b) Volume? 1

c) If you have a school bag of size 18.5 cm x14.5cm x 40cm, how many of such text books can you carry in it? 1

Appendix 2 (Contd..3)

IV Multiple choice questions. Circle your best answer. Marks 1x 8 = 8

1) Total surface area of a hemisphere is

$a)\ \pi r^2$h (b) $2\pi r^2$ (c) $3\pi r^2$ (d) 2/3πr^3

2) Which geometrical figure of the following has the slant height?

a) Cube (b) Cone (c) Circle (d) Square

3) Imran and Monika are given a cubical box each of the same size and are asked to fill up their individual boxes with sand from a nearby playground. Imran uses a conical flask and Monica uses a cylindrical can of the same height and base area. Who has to fill their containers with sand a minimum number of times and why, so that they should be able to complete the task faster?

a) Imran can do it quicker because he is stronger than Monica

b) Monica can do it quicker because she can run faster than Imran

c) Imran can do it quicker because he uses a container that is 3 times bigger than what Monica uses.

d) Monica can do it quicker as her container being cylindrical can carries much larger amount of sand.

4) A rubric (cuboid) occupies 25 sq.cm of base area. How much space can it occupy in air in cm^3?

a) 5 b) 5 c) 125 d) 25^3

5) A sports girl is practicing a marathon race of 3.140 kms in a circular track of radius 100 m. How many rounds would she need to make in order to complete the race?

a) Ten b) Five c) Twenty d) Fifty

6) A hemi-spherical bowl has a depth of 9 cms. How much milk can be filled into it completely, but without over-flowing? Your answer in terms of π is.

a) 486 π b) 972 π c) 162 π d) 324 π

APPENDIX 2 (Contd..3)

---

7) A watermelon, in the shape of a sphere measuring 20 cms in diameter, has been cut into two equal halves for ease of sale. For health reasons each cut piece needs to be wrapped completely and tightly for display before the customer delivery. The area of wrapping paper required for each piece in terms of π is:

a) 200π b) 300π c) 400π d) 1600π

8) An ice crème cone measures 5 cms in radius and 12 cms long. It is wrapped up completely and stored in a fridge before serving to customers. The wrapping paper is printed in colour attractively with company name and details at a cost of Rs.0.50 per sq.cm. The total printing cost of the paper just needed for the sale of 100 similar cones rounded to the nearest rupee would be (Use $\pi = 22/7$):

a) 7071 b) 14143 c) 3929 d) 2043

Note: The marking scheme for the Prior Knowledge test (pretest) is given below for each task type

---

Task types	Q.Nos	Max Score (25)
D1:	I (1) to I (8)	4
D2:	III (a), (b); IV (1), (2) and (3)	5
P1:	II (1) to II (4) and III (c)	11
P2:	IV (4) to IV (8)	5

Thank You

Appendix 3

TEACHERS' REFLECTIVE JOURNAL: A TEMPLATE

Objective: To be able to use students' prior knowledge for teaching a mathematics topic of year 10: Measurement

Date/time: **Teacher's name:**

I Learning Intention:

II Brief description of teaching task(s) used in the class:

III What questions were raised to invite discussion/participation of:

a) low ability students and who were given opportunity to answer?

b) Medium ability students, who answered and how?

c) High ability students, who answered and how?

IV a) How teacher has encouraged students to ask questions and facilitated learning?

b) Any other observations of students which help to improve their learning

---

Appendix 4

POSTTEST ON MEASUREMENT FOR CLASS 1: RESOURCE MATERIAL FOR MATHEMATICS SELF-EFFICACY SCALE (MSES)

Part A Max time allowed:45min

1. Change the units in the following as indicated

a) 8 m = _____ cm

b) 6 km = _____ m

c) 7 cm = _____ mm

d) 9 km = _____ cm

2. Complete the following:

a) 70 mm^2 = _____ cm^2

b) 2.3 m^2 = _____ cm^2

c) 0.004 km^2 = _____ m^2

d) 8.1 km^2 = _____ cm^2

3. Complete the following:

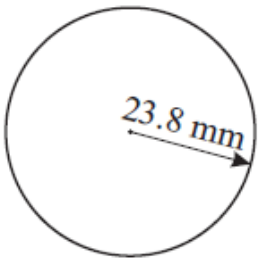
a) 7 cm^3 = _____ mm^3

b) 1.8 m^3 = _____ cm

c) 0.06 km^3 = _____ cm^3

d) 810 cm^3 = _____ km^3

4 Calculate the circumference of:

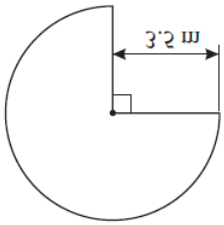


Appendix 4 (contd..2**)**

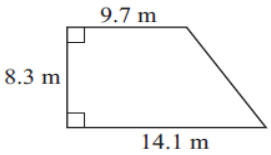
---

5. Calculate the PERIMETER of the following:

a)

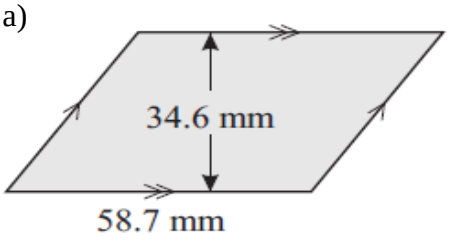


6

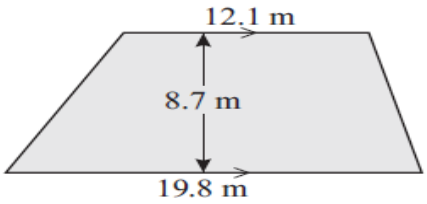


9 Calculate the shaded AREA of the following:

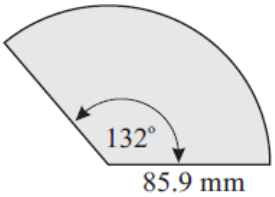
a)



b)



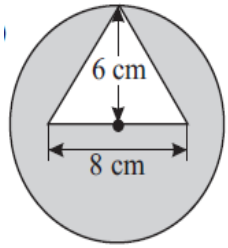
c)



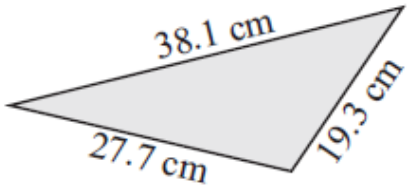
Appendix 4 (contd..3**)**

---

d) Calculate the shaded AREA of the following

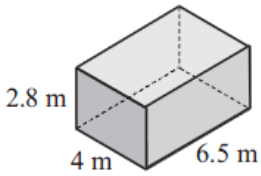


8 Use HERON'S FORMULA to calculate the area of:

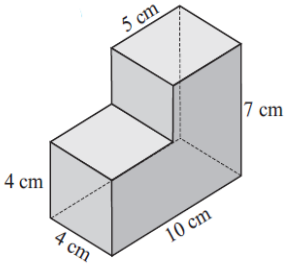


9 Calculate the TOTAL SURFACE AREA of the following:

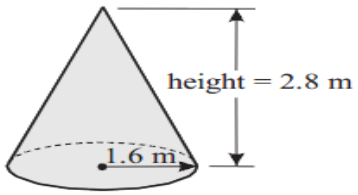
a)



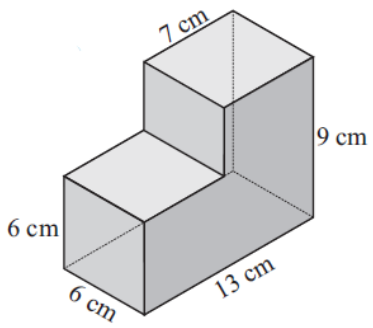
b)



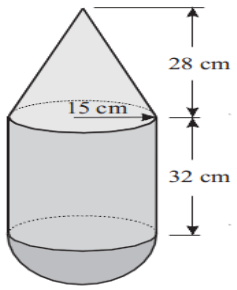
c) Calculate the TOTAL SURFACE AREA of the following:



d) Calculate the TOTAL SURFACE AREA of the following

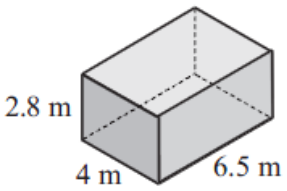


f)



10 Calculate the VOLUME of the following

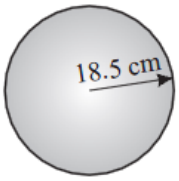
a)



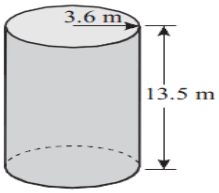
Appendix 4 (contd..5)

---

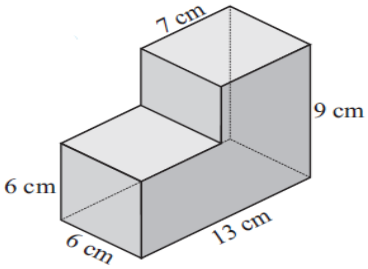
b) Calculate the VOLUME of the following



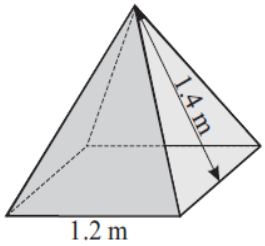
c)



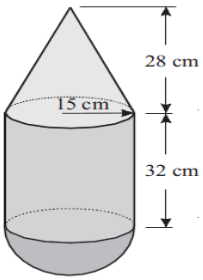
d)



e)



f)



THANK YOU

Appendix 5

MATHEMTICS SELF-EFFICACY SCALE (MSES) FOR CLASS 1

Directions: This questionnaire contains 3 pages of questions that correspond with question numbers of the resource material for the topic of measurement. Do **not** solve any question at this stage. Just go over each question quickly and select a response on a scale of 1 to 5. Your choice must indicate whether you can do similar type of question. This should indicate **your gut feeling only** (termed *self-efficacy belief* in successfully achieving or answering that question. There is **no correct or incorrect** choice when you respond. Your responses will be confidential.

Your responses must represent how surely you can answer each question from the resource material attached. The scores that you could expect for any question corresponding to each response are given below.

Response		Expected Score
1	if you **don't know** how to solve that problem	0 %
2	if you're **unsure**, but can try	25%
3	if you **think 'Maybe'**	50%
4	if you're **sure** you can solve the problem	75%
5	if you're **very sure** that you can do it successfully	100%

	Certainty rating				
Question number from Resource material	Don't know	Unsure	May be	Sure	Very sure
1.a	1	2	3	4	5
b).	1	2	3	4	5
Etc.					

---THANK YOU---

Appendix 6

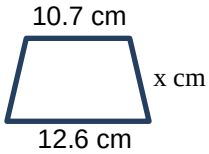
POSTTEST ON MEASUREMENT FOR CLASSES 3-16: RESOURCE MATERIAL FOR MATHEMATICS SELF-EFFICACY SCALE (MSES)

Part A Max time allowed:45min

Qn I Multiple choice questions

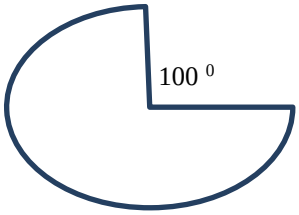
1. If the perimeter of this shape is 60.3 cm, the value of *x* is:
   A 12.8

   B 6.5
   C 5.7
   D 18.5
   E 3.5



2. For the sector of radius 5 m as shown below, the perimeter rounded to one decimal place is:

   A 8.7 m
   B 22.7 m
   C 18.7 m
   D 56.7 m
   E 32.7 m



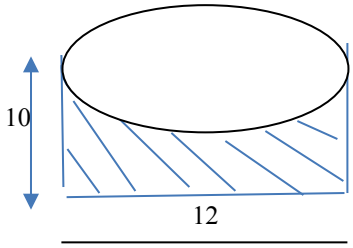
3. The exact value of the shaded area in the figure below in terms of π

   A 32 + 72π

   B 120 –36π

   C 32 + 6π

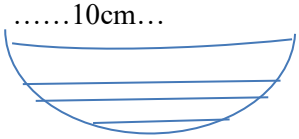
   D 120 –18π
   E 48 + 18



. 4. 0.128 m^2 is equivalent to:

   A 12.8 cm^2
   B 128 mm^2
   C 1280 cm^2
   D 0.00128 cm^2
   E 1280 mm^2

5. The volume of the hemispherical bowl shown below is:
   A 1257 cm^3 B 628 cm^3
   C 2094 cm^3 D. 1571 cm^3
   E.4189 cm^3



Part B

1.Change the units in the following as indicated low**:**

i. 8 m = _____ cm

ii. 6 km = _____ m

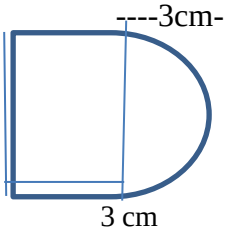
iii. 70 mm^2=_____ cm^2

iv. 2.3 m^2=_____ cm^2

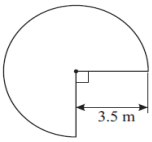
v. 0.004 km^2= _____ m^2

vi. 1.8 m^3 =_____ cm^3

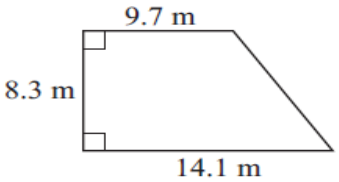
2. Calculate the total length of wire required to fence the paddock, which is a composite shape of a rectangle and a half-circle as shown below: Show working



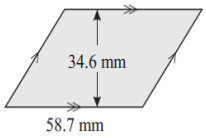
3a) Calculate the PERIMETER of the following shapes:



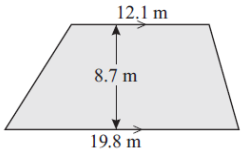
b) Perimeter of the deck Hint: First you will need to find the slant height using the Pythagoras theorem



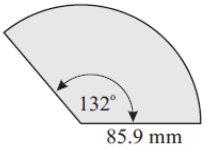
4. Calculate the shaded AREA of the following shapes:



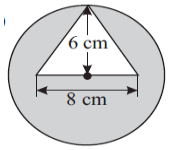
b)



c)

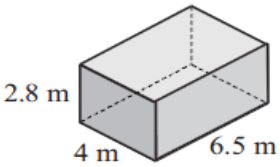


d)

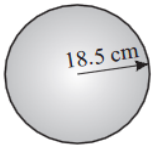


5. Work out the TOTAL SURFACE AREA of the following 3D shapes:

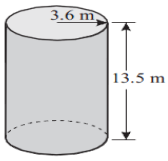
a)



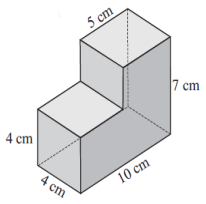
b) A shotput of radius 18.5cm



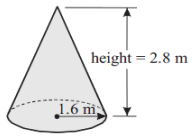
c) A cylinder of radius 3.6m and height 13.5m



d**)** Work out the TOTAL SURFACE AREA of the following 3D shapes

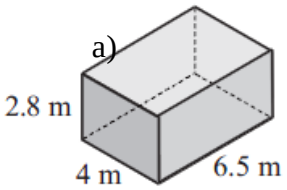


e)

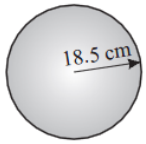


6.Calculate the VOLUME of each of the following 3D shapes**.**

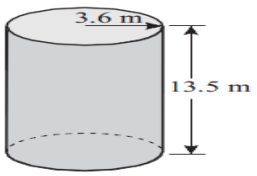
a)



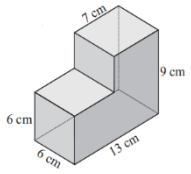
b) A globe



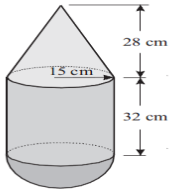
c)



d) Calculate the VOLUME of each of the following 3D shapes**:**



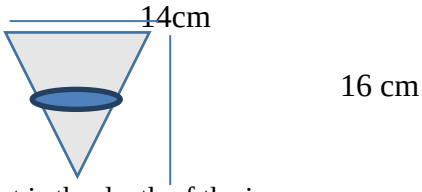
e)



Qn.7 An L-shaped concrete slab is required for the foundation of a new house. It is made up of two rectangles with dimensions 3 m by 2 m and 10 m by 6 m.

If two bags of cement are required for every 5 $m^2$of concrete, how many whole bags of cement will need to be purchased for the job? Show your working below.

Qn.8 A cone measuring 14 cm in diameter and 16 cm height is filled up to half of its volume with ice-crème.



a) What is the depth of the ice-cream correct to two decimal places?
(Hint: *Height(h) and radius (r) are in a constant ratio, that is, h/r* = k, a constant)
An ice crème lover has cut off the empty part of the cone exactly up to the ice-crème level.

b) Name the shape that has been separated from the cone and work out the area of unused paper to the nearest square centimetre.

Appendix 7

MATHEMTICS SELF-EFFICACY SCALE (MSES) FOR CLASSES 3-16

Directions: This questionnaire contains 3 pages of questions that correspond with question numbers of the resource material for the topic of measurement. Do **not** solve any question at this stage. Just go over each question quickly and select a response on a scale of 1 to 5. Your choice must indicate whether you can do similar type of question. This should indicate **your gut feeling only** (termed *self-efficacy belief* in successfully achieving or answering that question. There is **no correct or incorrect** choice when you respond. Your responses will be confidential.

Your responses must represent how surely you can answer each question from the resource material attached. The scores that you could expect for any question corresponding to each response are given below.

Response		Expected Score
1	if you **don't know** how to solve that problem	0 %
2	if you're **unsure**, but can try	25%
3	if you **think 'Maybe'**	50%
4	if you're **sure** you can solve the problem	75%
5	if you're **very sure** that you can do it successfully	100%

	Certainty rating				
Question number from Resource material	Don't know	Unsure	May be	Sure	Very sure
1.	1	2	3	4	5
2.	1	2	3	4	5
Etc.					

---THANK YOU---

Appendix 8

MATHEMATICS-RELATED CONSTRUCTIVIST ORIENTED LEARNING ENVIRONMENT SURVEY (MCOLES)

Name of student: ……. Gender----------

Directions

This section of the questionnaire contains statements about practices that could take place in your class. To describe how often each practice actually takes place in your class, choose a number

1	if this happens **Almost Never**
2	if this happens **Rarely**
3	if this happens **Sometimes**
4	if this happens **Often**
5	if this happens **Almost Always**

Student Cohesiveness and Personal Relevance

Item	Statement
D1	I make friends with many students in my mathematics class and many of them are already my friends.
D2	I know other students in my mathematics class and I work well with them.
D3	Students in my mathematics class like me because I am friendly with them.
D4	I help other class members who are having trouble with their mathematics work, and they help me too.
D5	I relate what I learn in my mathematics class to my life outside school and connect it.
D6	I draw on my past experiences and apply them to the work in my mathematics class.
D7	What I learn in my mathematics class is relevant to my everyday life in my school and outside.
D8	My mathematics class is relevant to my life because I get an understanding of life even outside of school.

Teacher Support

Item	Statement
D9	My mathematics teacher is interested in my mathematics problems.
D10	My mathematics teacher goes out of his/her way to help me.

Item	Statement
D11	My mathematics teacher considers my feelings.
D12	My teacher helps me when I have trouble with my mathematics work.
D13	My mathematics teacher talks with me about mathematics work.
D14	My mathematics teacher takes an interest in my progress.
D15	My mathematics teacher moves about the class to talk with me.
D16	My mathematics teacher's questions help me to understand.

Involvement

Item	Statement
D17	I discuss ideas in my mathematics class.
D18	I give my opinions during mathematics class discussions.
D19	My mathematics teacher asks me questions.
D20	I contribute to mathematics discussions in my class with my ideas and suggestions.
D21	I ask my mathematics teacher questions
D22	I explain my mathematics ideas to my peers.
D23	Students discuss with me how to go about solving problems.
D24	I am asked to explain how I solve problems.

Task orientation by cooperation

Item	Statement
D25	I cooperate with other students and learn from them when doing mathematics assignment work in the class.
D26	I share my mathematics books and resources with other students and cooperate with them when doing mathematics assignments in mathematics class.

Appendix 8 (contd..3)

Item	Statement
D27	When I work with others in groups in mathematics class, we work as a team to achieve class goals.
D28	I work on mathematics tasks with other students in my class.
D29	I know getting a certain amount of mathematics work done is important and also how much of mathematics work I have to do.
D30	I try to understand the mathematics work that I am required to do when completing a mathematics task.
D31	I know the goals set for my mathematics class.
D32	I am ready to pay attention to my mathematics teacher from the beginning until the end of the class.

Equity

Item	Statement
D33	My mathematics teacher gives me as much attention as to other students in my mathematics class.
D34	My mathematics teacher helps me as much as he does to others in my mathematics class.
D35	I have the same amount of say in my mathematics class as other students.
D36	I am treated the same as other students in my mathematics class.
D37	I receive the same encouragement from my mathematics teacher as other students do.
D38	I get the same opportunity to contribute to mathematics class discussions as other students.
D39	My mathematics work receives as much praise as other students' work.
D40	I get the same opportunity to answer mathematics questions as other students.

Differentiation

Item	Statement
D41	I work at the speed which suits my mathematics ability.
D42	Students who work faster than me can move on to the next mathematics topic.

Appendix 8 (contd..4)

Item	Statement
D43	I choose mathematics tasks suited to my interest.
D44	The mathematics tasks that are used in my class are suited to my interest.
D45	Mathematics tasks are suited to my ability.
D46	I use different mathematics materials from those used by other students.
D47	I use different mathematics assessment methods from other students.
D48	I do mathematics work that is different from other students' work.

Clarity of Assessment Criteria and Feedback

Item	Statement
D49	For improving my mathematics learning I use feedback from assessment tasks and understand their link with classroom activities.
D50	Mathematics assessment tasks are an important part of my learning as they help me to recognise my weaknesses in mathematics understanding.
D51	Mathematics assessment tasks help me to understand the topic.
D52	I find the mathematics assessment tasks meaningful and helpful to monitor my own learning.
D53	The criteria for mathematics assessment are clear to me as they inform me which activities and tasks are used to assess my performance.
D54	The requirements for assessment tasks are clear to me and I know what types of information I need for completing such tasks.
D55	I understand how my teacher judges my work from my teacher's instructions for doing assessment tasks.
D56	I know how to complete different assessment tasks successfully.

Note: This questionnaire is based on the Constructivist Oriented Classroom Learning Environment Survey (COLES), authored by Aldridge, Fraser, Bell and Dorman (2012). It was modified, used in my study and reproduced in this thesis with the permission of the authors.

Thank you

Appendix 9

WEIGHTS ASSIGNED TO TASKS ON POSTTEST

Table A.9.1: Weights (Scores) Assigned to Tasks on Posttest for Classes 1 and 3 - 16.

Task code	Class 1		Other Classes (3 - 16)	
	Q. No	Weight	Q. No	Weight
D11	7a	1.0	4a	1
D12	7b	1.0	4b	1
D13	7c	1.0	4c	1
D14	7d	2.0	4d	1
D21	1a	0.5	I(i)	1
D22	1b	0.5	I(ii)	1
D23	1c	0.5	I(iii)	1
D24	1d	0.5	I(iv)	1
D25	2a	0.5	I(v)	1
D26	2b	0.5	I(vi)	1
P101	9d	2.0	5d	2
P102	9e	1.0	5e	2
P103	10a	1.0	6a	2
P104	10b	1.0	6b	2
P105	10c	1.0	6c	2
P201	2c	0.5	2	2
P202	2d	0.5	3a	2
P203	3a	0.5	3b	2
P204	8	2.0	7	4
P205	10e	1.5	8a	2
P206	3b	0.5	8b	3
P301	9a	1.0	5a	2
P302	9b	1.0	5b	2
P303	9c	1.0	5c	2
P304	10d	2.0	6d	2
P305	10f	1.5	6f	2
P401	4	1.0	A1	1
P402	5	2.0	A2	1
P403	6	2.0	A3	1
P404	3c	0.5	A4	1
P405	3d	0.5	A5	1

Appendix 9 (contd..2)

Table A. 9.2: Weights (Scores) Assigned to Task Types: A Comparison between Classes

Task Type	Class 1	Other Classes	Task Types	Class 1	Other Classes
Declarative Knowledge Types			Procedural knowledge Types		
D1	15.6	8.0	P1	18.8	20.0
D2	9.4	12.0	P2	17.2	20.0
			P3	20.2	30.0
			P4	18.8	10.0
Total	25.0	20.0	Total	75.0	80.0

D1: Knowledge of facts and meanings
D2: Knowledge of concepts and definitions
P1: Integration of knowledge
P2: Application of knowledge
P3: Mixed tasks of integration of ideas and applications
P4: Challenging application tasks

Appendix 10

RESULTS OF INTERNAL CONSISTENCY, DISCRIMINANT VALIDITY, ANOVA AND EIGENVALUES FOR BOYS AND GIRLS

Table A10.1: Boys: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=313)

Scale	Unit of Analysis	No. of Items	Cronbach Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Student Cohesiveness & Personal Relevance after dropping item D4@	Student	7	0.629	0.120	0.100*
	Class mean	7	0.843		
Teacher Support after dropping item D15@	Student	7	0.842	0.120	0.170*
	Class mean	7	0.921		
Involvement after dropping item D19@	Student	7	0.778	0.120	0.080*
	Class mean	7	0.870		
Task Orientation by Co-operation				0.170	0.110*
Co-operation (D25, D26, D28)	Student	3	0.607		
	Class mean	3	0.722		
Task Orientation (D27, D29 to D32)	Student	5	0.638		
	Class mean	5	0.730		

* $p < 0.01$. The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students. @ D4, D15 and D19 were weakly correlated with other items, being less than 0.4

(contd..)

Appendix 10 (Contd..2)

Table A10.1 (contd..2): Boys: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=313)

Scale	Unit of Analysis	No. of Items	Cronbach Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Equity	Student	8	0.841	0.150	0.150*
	Class mean	8	0.940		
Differentiation (after dropping D42, D43) @				0.090	0.100*
Task Differentiation	Student	3	0.573		
(D41, D44, D45)	Class mean	3	0.651		
Peer Differentiation	Student	3	0.715		
(D46 to D48)	Class mean	3	0.795		
Clarity of Assessment	Student	7	0.791	0.100	0.110*
Criteria & Feedback	Class mean	7	0.915		
after dropping D49@					

* $p < 0.01$.The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students.

@ D42, D43 and D49 were weakly correlated with other items, being less than 0.4.

Appendix 10 (Contd..3)

Table A10.2: Girls: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=198)

Scale	Unit of Analysis	No. of Items	Cronbach Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Student Cohesiveness & Personal Relevance by dropping item D4@	Student	7	0.772	0.090	0.280*
	Class mean	7	0.805		
Teacher Support by dropping item D15@	Student	7	0.853	0.100	0.350*
	Class mean	7	0.953		
Involvement by dropping item D19@	Student	7	0.828	0.130	0.340*
	Class mean	7	0.942		
Task Orientation by Cooperation				0.150	
Co-operation (D25, D26, D28)	Student	3	0.763		
	Class mean	3	0.806		
Task Orientation (D27, D29 to D32)	Student	5	0.614		
	Class mean	5	0.835		

* $p < 0.01$. The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students

@ D4, D15 and D19 were weakly correlated with other items, being less than 0.4.

(contd...)

Appendix 10 (Contd..4)

Table A10.2 (contd..2): Girls: Internal Consistency Reliability (Cronbach's Alpha Coefficient), Discriminant Validity (Mean Partial Correlation with Other Scales) and Ability to Differentiate between Classes (eta^2 ratio from ANOVA) for Individual Student and Class Mean as Units of Analysis (N=198)

Scale	Unit of Analysis	No. of Items	Cronbach Alpha Coefficient	Mean Partial Correlation with Other Scales	eta^2 from ANOVA
Equity	Student	8	0.875	0.130	0.300*
	Class mean	8	0.929		
Differentiation (after dropping D42, D43) @				0.160	
Task Differentiation (D41, D44, D45)	Student	3	0.664		
	Class mean	3	0.705		
Peer Differentiation (D46 to D48)	Student	3	0.769		
	Class mean	3	0.917		
Clarity of Assessment Criteria & Feedback after dropping D49 @	Student	7	0.799	0.160	0.300*
	Class mean	7	0.921		

* $p < 0.01$. The eta^2 statistic is the ratio of 'between' to 'total' sums of squares, and it represents the proportion of variance explained by the class membership of students

@ D42, D43 and D49 were weakly correlated with other items, being less than 0.4

Appendix 10 (Contd..5)

Table A10.3: Boys: Eigenvalues and Factor Determination for MCOLES (N=313)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Student Cohesiveness & Personal Relevance (d=7, by dropping D4)	2.188 (31.3%) 1.435 (20.5%)	1.208 1.125	2	0.694	268.0	21
Teacher Support (d=7, by dropping D15)	3.558 (50.8%)	1.239	1	0.871	694.6	21
Involvement (d=7 by dropping D19)	3.073 (43.9%)	1.239	1	0.822	516.2	21
Task Orientation by Co-operation (d=8)	2.771 (34.6%) 1.268 (15.9%)	1.191 1.095	2	0.756	428.0	28

contd..)

Appendix 10 (Contd..6)

Table A10.3 (contd..2): Boys: Eigenvalues and Factor Determination for MCOLES (N=313)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Equity (d=8)	3.805 (47.6%)		1	0.870	787.6	28
Differentiation (d=6, by dropping D42, D43)	2.270 (37.8%) 1.281 (21.3%)	1.191 1.095	2	0.664	316.2	15
Clarity of Assessment Criteria & Feedback (d=7, by dropping D49)**	3.128 (44.7%)	1.208	1	0.844	515.0	21

Notes: For obtaining the eigenvalue factors, Principal Axis Factoring and Direct Oblimin rotation were used in all cases except when more than one factor was extracted. In that two-factor case, Varimax rotation was used. Figures within brackets are the respective % of variance explained by items used in extracting a factor.

* Bartlett test of sphericity was found to be significant using chi sq. at the 5% level for all scales.

** For Clarity of Assessment Criteria & Feedback, the initial solution with d=8 items were improved when the low correlated item D49 was dropped, which resulted in an increase in the % of variance explained by that factor for d=7.

Appendix 10 (Contd..7)

Table A10.4: Girls: Eigenvalues and Factor Determination for MCOLES (N=198)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Student Cohesiveness & Personal Relevance (d=7, by dropping D4)	2.980 (42.6%)		1	0.802	302.0	21
Teacher Support (d=7, by dropping D15)	4.090 (58.4%)		1	0.868	656.3	21
Involvement (d=7 by dropping D19)	3.493 (49.9%)		1	0.849	459.5	21
Task Orientation by Co-operation (d=8)	2.132 (26.6%) 2.095 (26.2%)	1.296 1.188	2	0.625	384.8	28

Notes: For obtaining the eigenvalue factors, Principal Axis Factoring and Direct Oblimin rotation were used in all cases except when more than one factor was extracted. In that two-factor case, Varimax rotation was used. Figures within brackets are the respective % of variance explained by items used in extracting a factor.

* Bartlett test of sphericity was found to be significant using chi-square at the 5% level for all scales.

(contd…)

Appendix 10 (Contd..8)

Table A10.4 (contd..2): Girls: Eigenvalues and Factor Determination for MCOLES (N= 198)

Scale (d=No.of items)	Eigenvalues Using Original Data (% of Var. Explained)	Eigenvalues by Monte Carlo Parallel Analysis	No. of Factors from Scree Plot	KMO Index of Sampling Adequacy	Bartlett test of Sphericity	
					Chi-sq.*	df
Equity (d=8)	4.307 (53.8%)		1	0.872	678.6	28
Differentiation (d=6, by dropping D42, D43)	2.297 (38.3%) 1.697 (28.3%)	1.251 1.130	2	0.644	312.8	15
Clarity of Assessment Criteria & Feedback (d=7, by dropping D49)	3.241(46.3%) 1.145 (16.4%)**	1.287 1.158	2	0.821	404.3	21

Notes: For obtaining the eigenvalue factors, Principal Axis Factoring and Direct Oblimin rotation were used in all cases except when more than one factor was extracted. In that two-factor case, Varimax rotation was used. Figures within brackets are the respective % of variance explained by items used in extracting a factor.

* Bartlett test of sphericity was found to be significant using chi square at the 5% level for all scales.

** For Clarity of Assessment Criteria & Feedback, the initial solution with d=8 items were improved when the low correlated Item D49 was dropped, which resulted in an increase in the % of variance explained by that factor alone for d=7. Two eigenvalues derived from the data set were greater than 1; however, the second value of 1.145 was less than the respective eigenvalue (1.158) obtained from the Monte Carlo PA, and so it was discarded.

Appendix 11

SCREE PLOTS FOR EIGENVALUE FACTORS OF MCOLES

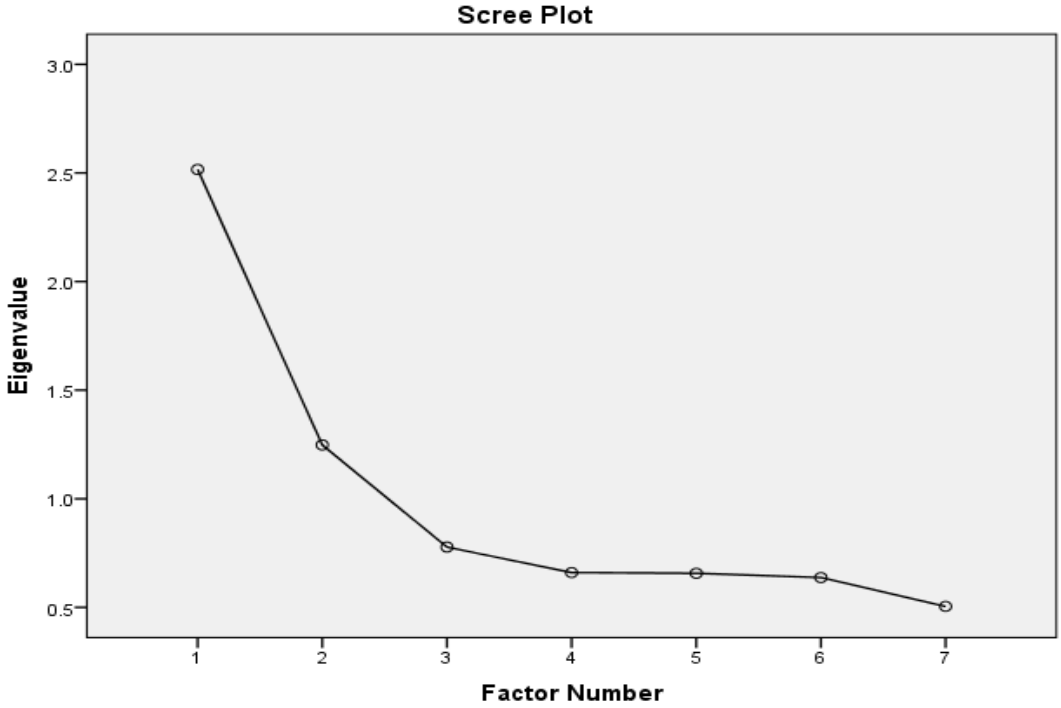


Figure A11.1: Scree Plot of Eigenvalue Factors for Student Cohesiveness and Personal Relevance: All Students (N=511)

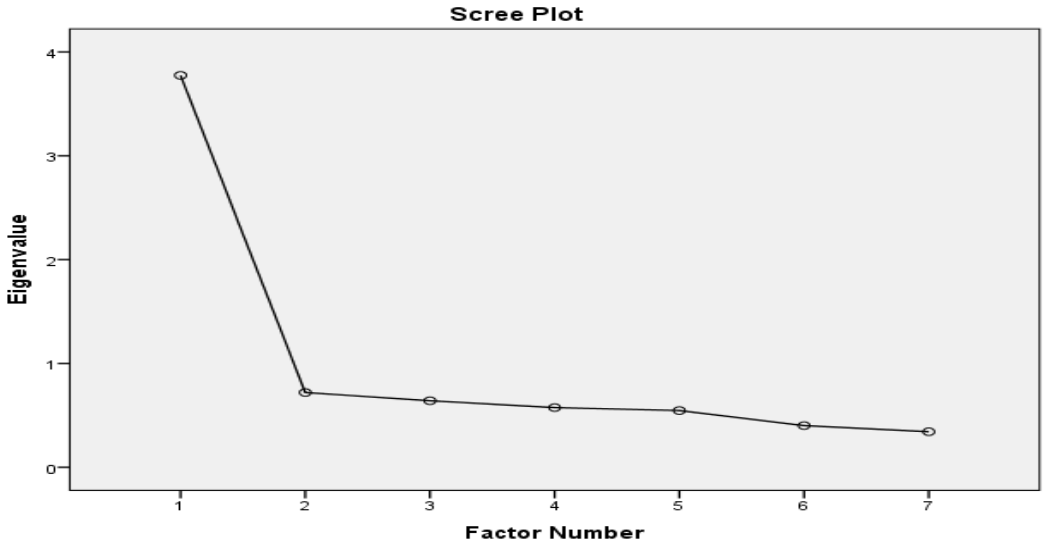


Figure A11.2: Scree Plot of Eigenvalue Factors for Teacher Support: All Students (N=511)

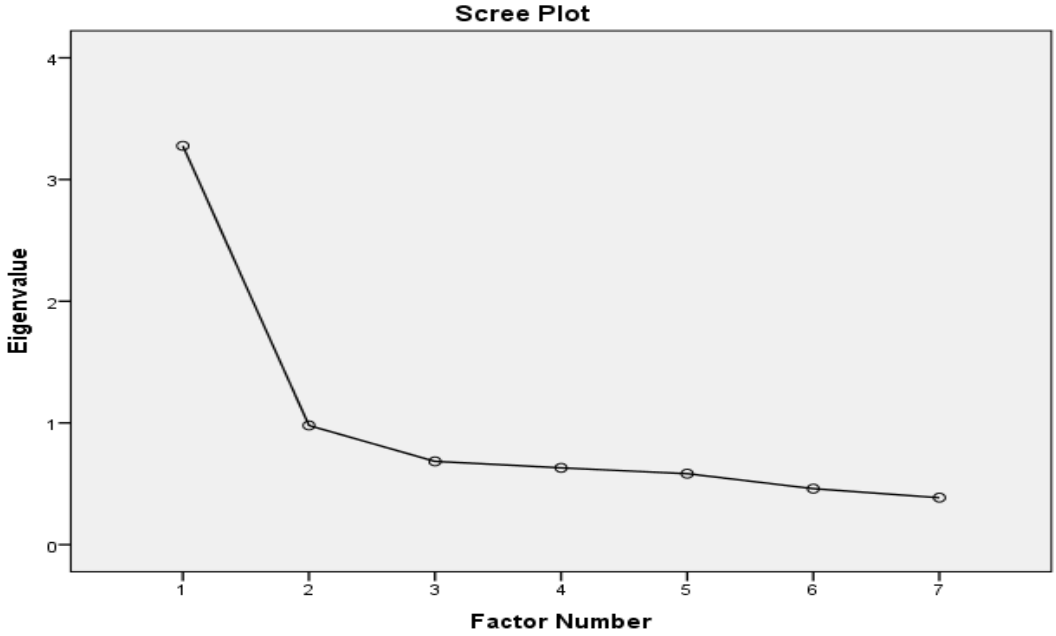


Figure A11.3: Scree Plot of Eigenvalue Factors for Involvement: All Students (N=511)

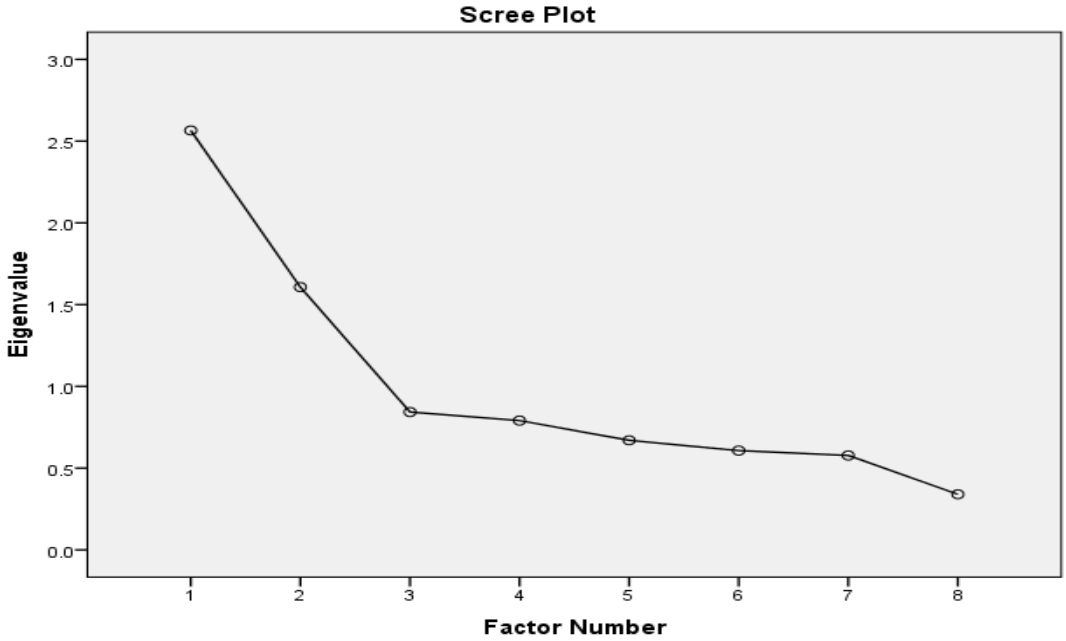


Figure A11.4: Scree Plot of Eigenvalue Factors for Task Orientation by Cooperation: All Students (N=511)

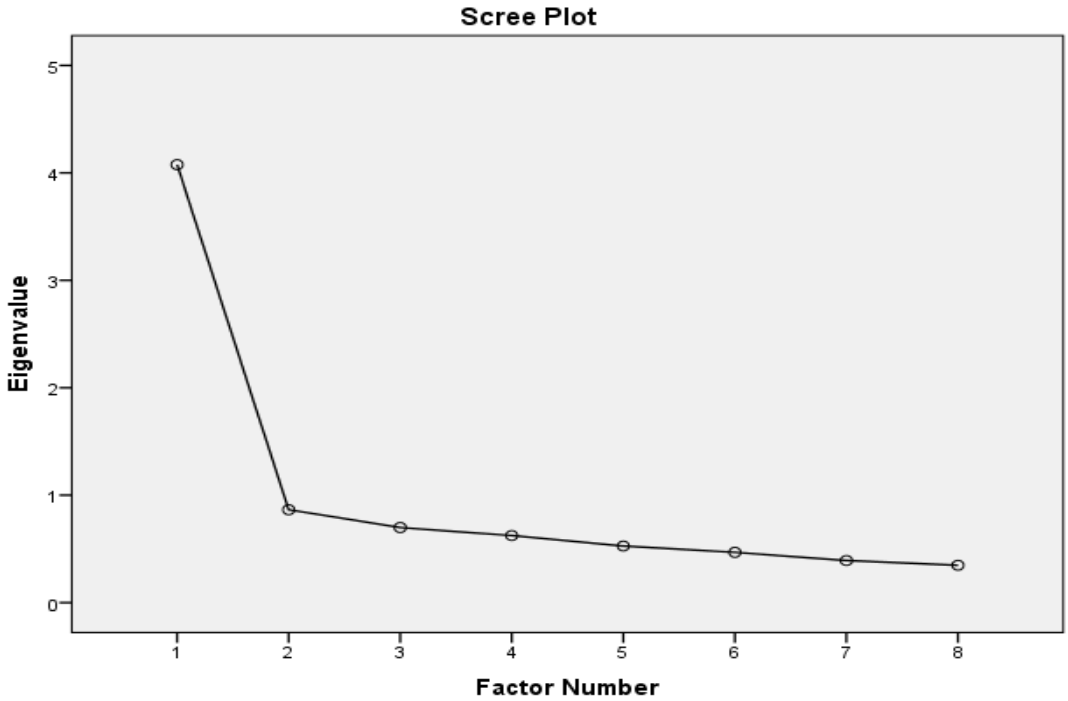


Figure A11.5: Scree Plot of Eigenvalue Factors for Equity: All Students (N=511)

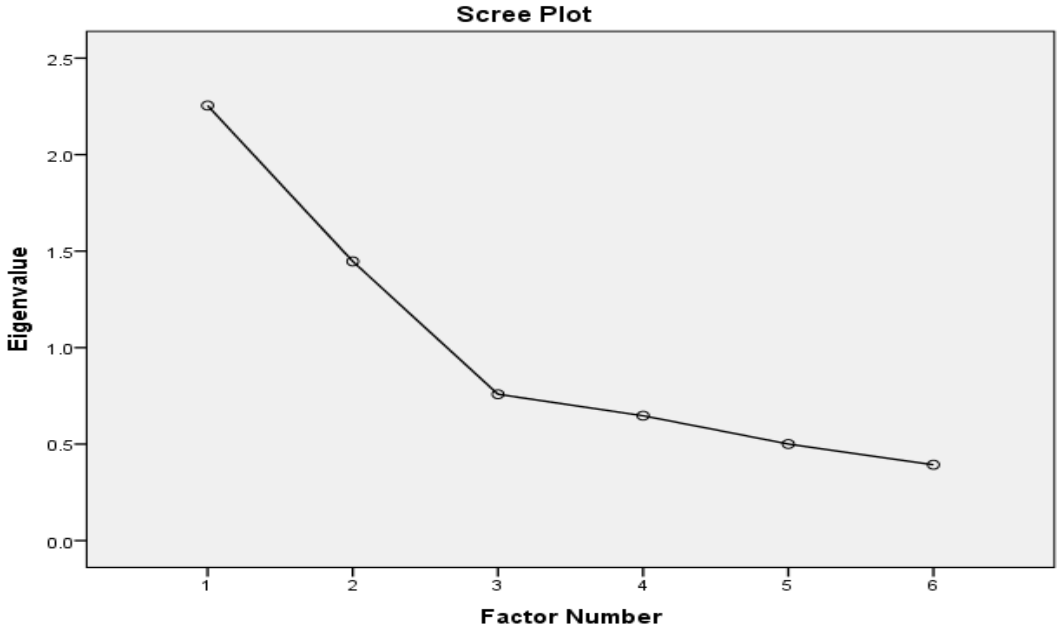


Figure A11.6: Scree Plot of Eigenvalue Factors for Differentiation: All Students (N=511)

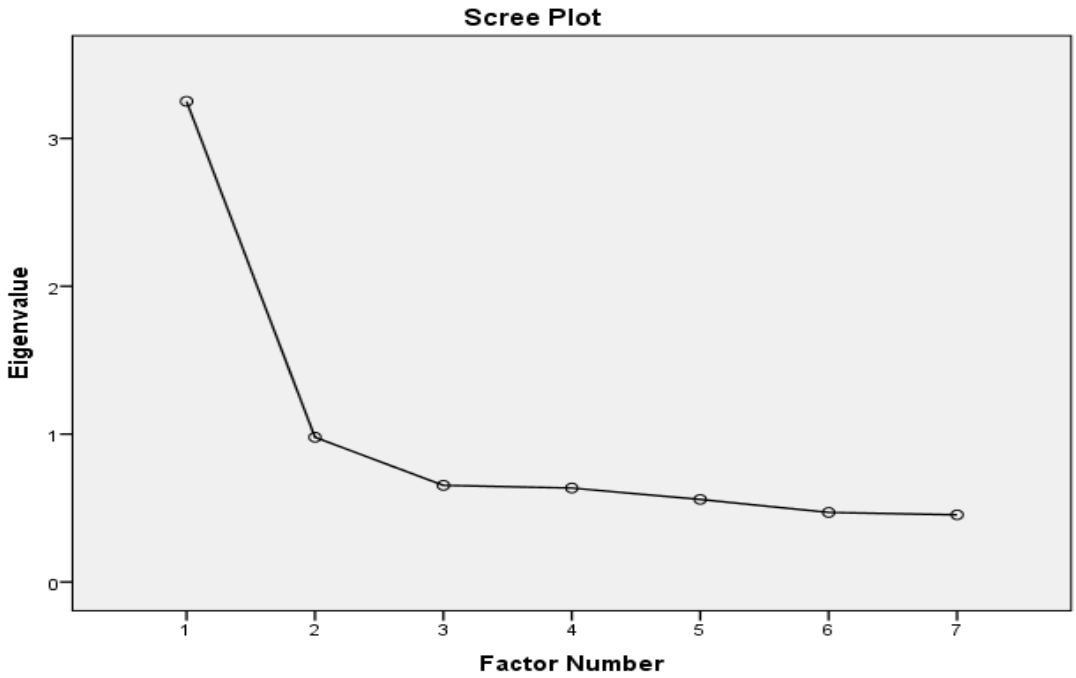


Figure A11.7: Scree Plot of Eigenvalue Factors for Clarity of Assessment Criteria & Feedback: All Students (N=511)

Appendix 12

LIST OF UNDERACHIEVERS IN EACH SCHOOL

Table A12.1: Underachievers in School 1

Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-Achievement %	Sex (Boys=0, Girls=1)
101002	1	50.0	49.6	23.0	115.7	0
101005	1	34.4	62.1	45.0	38.0	0
101006	1	37.5	32.3	30.0	7.5	1
101007	1	15.6	28.2	22.0	28.3	1
101009	1	62.5	80.6	57.0	41.5	0
101011	1	37.5	54.8	22.0	149.3	1
101013	1	43.8	75.8	70.0	8.3	0
101014	1	40.6	35.5	30.0	18.3	1
101017	1	43.8	80.6	70.0	15.2	1
101018	1	59.4	90.3	86.0	5.0	0
101019	1	18.8	29.0	8.0	262.9	1
101022	1	65.6	92.7	85.9	8.0	0

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.2: Underachievers in School 2

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-Achievement %	Sex (Boys=0, Girls=1)
1	203009	3	72.0	73.4	14.0	59.4	0
2	203030	3	68.0	70.4	18.0	52.4	0
3	203003	3	88.0	91.1	40.0	51.1	0
4	203016	3	70.0	66.1	16.0	50.1	0
5	203019	3	68.0	79.0	30.0	49.0	0
6	203001	3	52.0	71.0	24.0	47.0	0
7	203006	3	52.0	80.6	34.0	46.6	0
8	203031	3	88.0	79.8	34.0	45.8	0
9	203015	3	64.0	59.7	18.0	41.7	0
10	203008	3	60.0	73.6	32.0	41.6	0
11	203029	3	88.0	73.4	32.0	41.4	0
12	203017	3	88.0	84.7	44.0	40.7	0
13	203021	3	52.0	63.9	24.0	39.9	0
14	203032	3	84.0	62.2	28.0	34.2	0
15	203025	3	60.0	54.1	20.0	34.1	0
16	203013	3	68.0	58.9	28.0	30.9	0
17	203038	3	60.0	58.9	28.0	30.9	0
18	203026	3	72.0	56.5	28.0	28.5	0
19	203039	3	76.0	50.0	24.0	26.0	0
20	203023	3	52.0	41.9	16.0	25.9	0
21	203040	3	54.0	56.5	42.0	14.5	0
22	203033	3	72.0	94.4	26.0	68.4	1
23	203036	3	72.0	93.5	38.0	55.5	1
24	203037	3	72.0	74.2	24.0	50.2	1
25	203004	3	68.0	71.8	30.0	41.8	1
26	203002	3	56.0	84.6	44.0	40.6	1
27	203007	3	68.0	74.2	34.0	40.2	1
28	203020	3	72.0	62.9	24.0	38.9	1
29	203022	3	64.0	69.4	36.0	33.4	1
30	203034	3	90.0	67.7	40.0	27.7	1
31	203014	3	56.0	62.9	36.0	26.9	1
32	203011	3	90.0	65.3	42.0	23.3	1
33	203024	3	68.0	58.7	36.0	22.7	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.2 (contd..3): Underachievers in School 2

Sr. No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
34	203012	3	64.0	52.4	32.0	20.4	1
35	203010	3	60.0	49.2	32.0	17.2	1
36	203028	3	68.0	49.2	32.0	17.2	1
37	203018	3	56.0	49.2	42.0	7.2	1
38	203027	3	80.0	53.2	48.0	5.2	1
39	204022	4	63.6	79.0	12.0	67.0	0
40	204008	4	83.6	93.5	34.0	59.5	0
41	204012	4	83.2	97.6	46.0	51.6	0
42	204002	4	90.0	95.2	44.0	51.2	0
43	204013	4	83.6	60.9	10.0	50.9	0
44	204027	4	90.0	96.0	47.0	49.0	0
45	204005	4	82.3	54.7	10.0	44.7	0
46	204007	4	80.9	83.1	40.0	43.1	0
47	204025	4	67.0	78.2	38.0	40.2	0
48	204016	4	80.5	83.1	45.0	38.1	0
49	204038	4	85.5	100.0	62.0	38.0	0
50	204001	4	86.9	79.8	42.0	37.8	0
51	204026	4	87.7	95.6	60.0	35.6	0
52	204011	4	80.9	69.4	38.0	31.4	0
53	204003	4	56.4	61.3	30.0	31.3	0
54	204031	4	82.7	65.3	38.0	27.3	0
55	204043	4	92.7	98.4	72.0	26.4	0
56	204023	4	63.6	61.8	36.0	25.8	0
57	204041	4	80.9	84.7	64.0	20.7	0
58	204020	4	63.6	51.6	32.0	19.6	0
59	204029	4	73.2	81.6	62.0	19.6	0
60	204039	4	85.5	54.0	38.0	16.0	0
61	204015	4	85.5	54.8	40.0	14.8	0
62	204036	4	85.5	55.6	44.0	11.6	0
63	204019	4	56.4	47.6	38.0	9.6	0
64	204017	4	92.7	61.3	52.0	9.3	0
65	204018	4	56.4	48.0	40.0	8.0	0
66	204034	4	77.7	93.5	36.0	57.5	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.2 (Contd..4): Underachievers in School 2

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
67	204030	4	68.0	91.1	41.0	50.1	1
68	204006	4	44.1	78.0	29.0	49.0	1
69	204032	4	60.0	91.1	44.0	47.1	1
70	204004	4	56.4	79.8	40.0	39.8	1
71	204040	4	75.5	73.6	38.0	35.6	1
72	204035	4	80.5	75.8	46.0	29.8	1
73	204044	4	73.2	71.5	46.0	25.5	1
74	204014	4	63.6	73.4	49.0	24.4	1
75	205023	5	51.4	100.0	24.0	76.0	0
76	205025	5	46.4	96.8	32.0	64.8	0
77	205036	5	46.0	81.5	18.0	63.5	0
78	205029	5	43.4	80.6	22.0	58.6	0
79	205003	5	58.6	86.3	38.0	48.3	0
80	205028	5	48.0	87.1	44.0	43.1	0
81	205034	5	66.4	54.7	16.0	38.7	0
82	205004	5	64.1	91.9	56.0	35.9	0
83	205010	5	62.5	73.4	24.0	49.4	1
84	205033	5	72.5	74.2	26.0	48.2	1
85	205013	5	54.1	71.0	24.0	47.0	1
86	205017	5	64.1	62.9	26.0	36.9	1
87	205032	5	58.6	57.3	26.0	31.3	1
88	205021	5	66.4	59.7	40.0	19.7	1
89	205020	5	60.1	59.7	42.0	17.7	1
90	205035	5	68.0	59.7	44.0	15.7	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.3: Underachievers in School 3

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-Achievement %	Sex (Boys=0 Girls=1)
1	306031	6	92.0	83.1	38.0	45.1	0
2	306028	6	88.0	100.0	60.0	40.0	0
3	306003	6	84.0	87.1	60.0	27.1	0
4	306034	6	80.0	58.9	32.0	26.9	0
5	306017	6	68.0	85.5	62.0	23.5	0
6	306023	6	80.0	75.8	54.0	21.8	0
7	306038	6	68.0	83.1	62.0	21.1	0
8	306010	6	60.0	79.0	64.0	15.0	0
9	306019	6	80.0	89.5	76.0	13.5	0
10	306008	6	44.0	80.6	68.0	12.6	0
11	306014	6	64.0	46.0	36.0	10.0	0
12	306041	6	84.0	63.7	54.0	9.7	0
13	306018	6	84.0	79.4	72.0	7.4	0
14	306037	6	48.0	70.2	24.0	46.2	1
15	306005	6	60.0	91.1	52.0	39.1	1
16	306027	6	32.0	81.5	52.0	29.5	1
17	306013	6	92.0	80.6	52.0	28.6	1
18	306011	6	48.0	76.6	50.0	26.6	1
19	306025	6	28.0	71.8	52.0	19.8	1
20	306002	6	60.0	79.8	64.0	15.8	1
21	306033	6	64.0	73.4	58.0	15.4	1
22	306016	6	76.0	62.1	50.0	12.1	1
23	306036	6	72.0	62.1	56.0	6.1	1
24	306012	6	48.0	87.1	82.0	5.1	1
25	307011	7	80.0	71.8	16.0	55.8	0
26	307016	7	40.0	48.4	14.0	34.4	0
27	307009	7	80.0	73.4	48.0	25.4	0
28	307004	7	44.0	50.8	26.0	24.8	0
29	307010	7	56.0	54.8	32.0	22.8	0
30	307007	7	72.0	46.0	24.0	22.0	0
31	307002	7	56.0	66.1	52.0	14.1	0
32	307012	7	68.0	45.2	34.0	11.2	0
33	307003	7	68.0	58.1	48.0	10.1	0

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.3 (contd..2): Underachievers in School 3

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievemer %	Under-achievement %	Sex (Boys=0, Girls=1)
34	307008	7	76.0	71.8	42.0	29.8	1
35	307017	7	88.0	76.6	52.0	24.6	1
36	307005	7	72.0	64.5	58.0	6.5	1
37	307013	7	80.0	62.1	56.0	6.1	1
38	308002	8	100.0	96.0	68.0	28.0	0
39	308014	8	76.0	81.5	54.0	27.5	0
40	308011	8	76.0	87.9	74.0	13.9	0
41	308022	8	80.0	69.4	56.0	13.4	0
42	308005	8	100.0	94.4	86.0	8.4	0
43	308009	8	100.0	96.8	90.0	6.8	0
44	308028	8	72.0	85.5	80.0	5.5	0
45	308025	8	60.0	63.7	44.0	19.7	1
46	308016	8	68.0	89.5	70.0	19.5	1
47	308032	8	88.0	71.8	60.0	11.8	1
48	308035	8	68.0	71.0	60.0	11.0	1
49	308021	8	52.0	75.8	66.0	9.8	1
50	308030	8	92.0	89.5	80.0	9.5	1
51	308003	8	68.0	66.1	58.0	8.1	1
52	308013	8	72.0	75.0	68.0	7.0	1
53	308026	8	84.0	86.3	80.0	6.3	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.4: Underachievers in School 4

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
1	409021	9	72.0	82.3	38.0	44.3	0
2	409001	9	60.0	80.0	46.0	34.0	0
3	409020	9	80.0	83.1	58.0	25.1	0
4	409024	9	92.0	79.0	58.0	21.0	0
5	409036	9	48.0	60.2	44.0	16.2	0
6	409027	9	64.0	57.3	46.0	11.3	0
7	409014	9	84.0	81.2	70.0	11.2	0
8	409016	9	60.0	45.0	34.0	11.0	0
9	409019	9	60.0	44.9	34.0	10.9	0
10	409008	9	68.0	62.5	52.0	10.5	0
11	409034	9	76.0	79.8	70.0	9.8	0
12	409026	9	68.0	53.2	44.0	9.2	0
13	409039	9	84.0	66.1	58.0	8.1	0
14	409015	9	76.0	58.1	50.0	8.1	0
15	409012	9	72.0	70.2	62.0	8.2	1
16	409030	9	84.0	70.2	62.0	8.2	1
17	410010	10	24.0	59.7	48.0	11.7	0

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.5: Underachievers in School 5

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
1	511016	11	48.0	90.3	14.0	76.3	0
2	511026	11	36.0	85.5	50.0	35.5	0
3	511001	11	32.0	69.4	34.0	35.4	0
4	511020	11	56.0	75.0	40.0	35.0	0
5	511030	11	32.0	91.1	58.0	33.1	0
6	511035	11	40.0	83.1	50.0	33.1	0
7	511002	11	48.0	90.3	61.0	29.3	0
8	511023	11	52.0	83.1	55.0	28.1	0
9	511028	11	36.0	79.0	51.0	28.0	0
10	511014	11	28.0	76.6	50.0	26.6	0
11	511036	11	32.0	71.0	46.0	25.0	0
12	511003	11	56.0	83.9	59.0	24.9	0
13	511025	11	32.0	80.6	57.0	23.6	0
14	511019	11	68.0	97.6	74.0	23.6	0
15	511012	11	52.0	63.7	41.0	22.7	0
16	511017	11	56.0	90.3	69.0	21.3	0
17	511029	11	44.0	75.8	56.0	19.8	0
18	511006	11	48.0	63.7	45.0	18.7	0
19	511018	11	68.0	90.8	74.0	16.8	0
20	511015	11	44.0	87.9	51.0	36.9	1
21	511034	11	64.0	89.5	56.0	33.5	1
22	511027	11	52.0	88.7	57.0	31.7	1
23	511010	11	56.0	93.5	62.0	31.5	1
24	511005	11	68.0	91.1	61.0	30.1	1
25	511031	11	56.0	87.9	60.0	27.9	1
26	511009	11	56.0	63.7	36.0	27.7	1
27	511022	11	72.0	93.5	67.0	26.5	1
28	511033	11	64.0	86.3	62.0	24.3	1
29	511013	11	68.0	89.5	66.0	23.5	1
30	511008	11	60.0	83.9	61.0	22.9	1
31	511007	11	72.0	92.7	70.0	22.7	1
32	511021	11	64.0	88.7	66.0	22.7	1
33	511011	11	72.0	87.1	66.0	21.1	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.5 (contd..2): Underachievers in School 5

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
34	511004	11	64.0	87.1	67.0	20.1	1
35	512004	12	36.0	61.3	20.0	41.3	0
36	512024	12	20.0	66.1	26.0	40.1	0
37	512002	12	40.0	66.9	28.0	38.9	0
38	512028	12	32.0	64.5	26.0	38.5	0
39	512003	12	48.0	65.3	28.0	37.3	0
40	512013	12	44.0	77.4	42.0	35.4	0
41	512005	12	36.0	56.5	22.0	34.5	0
42	512009	12	20.0	61.3	28.0	33.3	0
43	512018	12	40.0	79.8	50.0	29.8	0
44	512029	12	56.0	89.5	62.0	27.5	0
45	512006	12	36.0	41.9	18.0	23.9	0
46	512007	12	40.0	66.9	44.0	22.9	0
47	512001	12	28.0	64.5	42.0	22.5	0
48	512016	12	28.0	64.5	20.0	44.5	1
49	512012	12	16.0	61.3	18.0	43.3	1
50	512020	12	16.0	67.7	26.0	41.7	1
51	512017	12	60.0	93.7	52.0	41.7	1
52	512023	12	48.0	65.3	30.0	35.3	1
53	512015	12	24.0	66.9	32.0	34.9	1
54	512025	12	24.0	46.8	12.0	34.8	1
55	512011	12	40.0	75.0	42.0	33.0	1
56	512019	12	28.0	73.4	42.0	31.4	1
57	512014	12	32.0	53.2	24.0	29.2	1
58	512022	12	32.0	64.5	36.0	28.5	1
59	512010	12	36.0	75.8	50.0	25.8	1
60	512021	12	52.0	85.5	60.0	25.5	1
61	512026	12	36.0	75.8	56.0	19.8	1
62	513011	13	48.0	74.2	36.0	38.2	0
63	513022	13	44.0	71.0	34.0	37.0	0
64	513017	13	28.0	56.5	20.0	36.5	0
65	513025	13	40.0	73.4	38.0	35.4	0
66	513036	13	40.0	75.8	42.0	33.8	0

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Appendix 12 (contd..10)

Table A12.5 (contd..3): Underachievers in School 5

Sr. No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
67	513002	13	48.0	72.6	40.0	32.6	0
68	513029	13	48.0	67.7	38.0	29.7	0
69	513024	13	36.0	45.2	16.0	29.2	0
70	513012	13	44.0	66.1	38.0	28.1	0
71	513020	13	60.0	79.8	52.0	27.8	0
72	513010	13	40.0	58.1	32.0	26.1	0
73	513032	13	32.0	72.6	48.0	24.6	0
74	513006	13	32.0	69.4	48.0	21.4	0
75	513021	13	48.0	75.0	56.0	19.0	0
76	513005	13	60.0	66.9	48.0	18.9	0
77	513001	13	64.0	70.2	58.0	12.2	0
78	513033	13	44.0	70.2	60.0	10.2	0
79	513018	13	56.0	58.9	32.0	26.9	1
80	513007	13	40.0	60.5	34.0	26.5	1
81	513027	13	56.0	78.2	52.0	26.2	1
82	513013	13	36.0	62.1	42.0	20.1	1
83	513026	13	48.0	72.6	54.0	18.6	1
84	513030	13	52.0	63.7	46.0	17.7	1
85	513003	13	44.0	76.6	60.0	16.6	1
86	513031	13	36.0	63.7	54.0	9.7	1
87	513009	13	52.0	84.7	78.0	6.7	1
88	513028	13	48.0	67.7	62.0	5.7	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.6: Underachievers in School 6

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
1	614010	14	96.0	100.0	76.0	24.0	0
2	614009	14	88.0	96.8	80.0	16.8	0
3	614029	14	84.0	96.0	80.0	16.0	0
4	614005	14	96.0	97.6	82.0	15.6	0
5	614023	14	96.0	96.8	84.0	12.8	0
6	614026	14	76.0	88.7	76.0	12.7	0
7	614013	14	88.0	96.0	84.0	12.0	0
8	614001	14	72.0	95.2	84.0	11.2	0
9	614003	14	64.0	87.1	78.0	9.1	0
10	614007	14	80.0	90.3	82.0	8.3	0
11	614016	14	84.0	91.9	84.0	7.9	0
12	614017	14	100.0	93.5	86.0	7.5	0
13	614028	14	80.0	89.5	82.0	7.5	0
14	614031	14	96.0	88.7	82.0	6.7	0
15	614006	14	64.0	84.7	78.0	6.7	0
16	614004	14	92.0	86.3	80.0	6.3	0
17	614014	14	76.0	78.2	72.0	6.2	0
18	614012	14	68.0	84.7	79.0	5.7	0
19	614027	14	64.0	80.6	75.0	5.6	0
20	614020	14	76.0	81.3	76.0	5.3	0
21	615003	15	64.0	75.0	50.0	25.0	0
22	615002	15	96.0	83.1	68.0	15.1	0
23	615004	15	96.0	93.5	82.0	11.5	0
24	615015	15	88.0	96.8	88.0	8.8	0
25	615014	15	92.0	86.3	78.0	8.3	0
26	615006	15	92.0	100.0	92.0	8.0	0
27	615009	15	80.0	87.9	80.0	7.9	0
28	615021	15	80.0	76.6	70.0	6.6	0
29	615024	15	88.0	98.5	92.0	6.5	0
30	616022	16	96.0	89.5	50.0	39.5	1
31	616028	16	96.0	88.8	52.0	36.8	1
32	616037	16	96.0	91.4	56.0	35.4	1
33	616009	16	96.0	90.3	56.0	34.3	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Table A12.6 (contd..2): Underachievers in School 6

Sr. No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Under-achievement %	Sex (Boys=0, Girls=1)
34	616023	16	96.0	91.1	62.0	29.1	1
35	616034	16	88.0	66.1	42.0	24.1	1
36	616010	16	100.0	96.0	72.0	24.0	1
37	616026	16	100.0	87.1	64.0	23.1	1
38	616017	16	100.0	96.8	74.0	22.8	1
39	616018	16	100.0	90.3	68.0	22.3	1
40	616001	16	88.0	79.8	58.0	21.8	1
41	616014	16	100.0	79.4	58.0	21.4	1
42	616033	16	100.0	88.7	68.0	20.7	1
43	616005	16	96.0	70.2	50.0	20.2	1
44	616013	16	92.0	85.5	66.0	19.5	1
45	616016	16	100.0	91.1	72.0	19.1	1
46	616004	16	100.0	88.7	70.0	18.7	1
47	616008	16	92.0	70.2	52.0	18.2	1
48	616011	16	92.0	70.2	52.0	18.2	1
49	616020	16	96.0	91.9	74.0	17.9	1
50	616025	16	80.0	89.5	72.0	17.5	1
51	616024	16	96.0	50.8	34.0	16.8	1
52	616015	16	92.0	76.0	62.0	14.0	1
53	616031	16	100.0	75.8	64.0	11.8	1
54	616021	16	100.0	66.9	56.0	10.9	1
55	616032	16	100.0	81.5	72.0	9.5	1
56	616012	16	96.0	81.5	74.0	7.5	1
57	616007	16	96.0	57.3	51.0	6.3	1
58	616027	16	100.0	73.4	68.0	5.4	1

Note: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively. Underachievement Index was measured by 100 x (Expected scores - Postscores) / Postscores). Students were treated as underachievers when the index value ranged above 5%.

Appendix 13

LIST OF ACHIEVERS WHO MATCHED SELF-EFFICACY EXPECTANCY

Table A13.1: Achievers Who Matched Self-Efficacy

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Deviation Index %	Sex (Boys=0 Girls=1)
1	101008	1	53.1	66.9	69.0	-3.0	1
2	101021	1	37.5	79.8	78.1	2.2	1
3	204009	4	63.6	38.1	40.0	-4.8	1
4	204037	4	90.5	62.0	64.0	-3.2	0
5	204042	4	77.3	36.5	36.0	1.3	0
6	306004	6	72.0	59.7	58.0	2.9	0
7	306015	6	96.0	75.0	72.0	4.2	0
8	306024	6	60.0	76.6	74.0	3.5	1
9	306029	6	76.0	57.3	60.0	-4.6	1
10	306030	6	96.0	93.5	92.0	1.7	0
11	306039	6	84.0	84.7	82.0	3.3	0
12	306040	6	96.0	66.1	66.0	0.2	0
13	307006	7	48.0	36.6	38.0	-3.6	0
14	307015	7	60.0	35.5	34.0	4.4	0
15	308004	8	64.0	68.5	68.0	0.8	1
16	308007	8	100.0	87.9	90.0	-2.3	0
17	308008	8	100.0	95.2	94.0	1.2	0
18	308015	8	92.0	93.5	92.0	1.7	0
19	308017	8	72.0	71.0	72.0	-1.4	1
20	308018	8	96.0	59.4	60.0	-0.9	0
21	308027	8	88.0	93.5	92.0	1.7	0
22	308033	8	88.0	87.9	92.0	-4.5	1
23	308039	8	72.0	85.5	86.0	-0.6	0
24	409005	9	88.0	85.5	86.0	-0.6	1
25	409018	9	100.0	86.3	88.0	-1.9	1
26	409022	9	88.0	56.5	58.0	-2.7	1
27	409023	9	76.0	56.5	56.0	0.8	0
28	409033	9	88.0	71.8	70.0	2.5	1
29	409037	9	48.0	42.8	42.0	1.8	0
30	410003	10	32.0	81.5	80.0	1.8	0
31	410008	10	48.0	51.8	52.0	-0.4	0
32	410014	10	64.0	75.0	74.0	1.4	0
33	410023	10	44.0	80.6	84.0	-4.0	1
34	410028	10	48.0	88.7	88.0	0.8	1
35	410031	10	28.0	78.2	78.0	0.3	0

Notes: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively.
Deviation Index was measured by 100 x (Expected scores - Postscores) / Postscores) in the range of ±5%.

Table A13.1 (Contd..2): Achievers Who Matched Self-Efficacy

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Deviation Index %	Sex (Boys=0, Girls=1)
36	410034	10	36.0	78.9	78.0	1.2	0
37	614002	14	48.0	79.8	78.0	2.4	0
38	614011	14	68.0	82.3	80.0	2.8	0
39	614018	14	80.0	75.0	74.0	1.4	0
40	614019	14	96.0	75.0	76.0	-1.3	0
41	614022	14	68.0	75.3	76.0	-0.9	0
42	614030	14	64.0	69.4	72.0	-3.7	0
43	615001	15	100.0	83.9	88.0	-4.7	0
44	615008	15	96.0	89.5	88.0	1.7	0
45	615010	15	16.0	75.8	76.0	-0.3	0
46	615013	15	92.0	89.9	90.0	-0.1	0
47	615016	15	100.0	92.7	92.0	0.8	0
48	615017	15	76.0	83.9	88.0	-4.7	0
49	615018	15	84.0	87.9	90.0	-2.3	0
50	615019	15	80.0	79.0	80.0	-1.2	0
51	615026	15	92.0	84.7	88.0	-3.8	0
52	615027	15	60.0	87.1	90.0	-3.2	0
53	615030	15	88.0	84.7	86.0	-1.5	0
54	616006	16	100.0	75.8	78.0	-2.8	1

Notes: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively.
Deviation Index was measured by 100 x (Expected scores - Postscores) / Postscores) in the range of ±5%.

Appendix 14

LIST OF ACHIEVERS WITH LOWER SELF-EFFICACY EXPECTANCY

Table A14.1: Achievers Having Lower Self-Efficacy

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Deviation Index %	Sex (Boys=0 Girls=1)
1	101004	1	25.0	50.0	73.0	-31.5	1
2	101010	1	90.6	39.5	77.0	-48.7	0
3	101012	1	46.9	33.9	55.0	-38.4	0
4	101015	1	78.1	66.1	84.0	-21.3	1
5	101016	1	25.0	46.0	64.0	-28.2	1
6	101020	1	31.3	53.2	77.0	-30.9	1
7	101024	1	71.9	78.2	84.4	-7.3	1
8	204033	4	82.7	50.0	62.0	-19.4	0
9	306001	6	72.0	63.1	74.0	-14.7	0
10	306006	6	96.0	67.7	72.0	-5.9	0
11	306021	6	76.0	75.8	86.0	-11.9	1
12	306022	6	72.0	73.4	80.0	-8.3	0
13	306026	6	56.0	71.8	86.0	-16.5	1
14	306035	6	68.0	72.6	82.0	-11.5	1
15	308001	8	68.0	67.7	100.0	-32.3	1
16	308006	8	100.0	91.1	100.0	-8.9	0
17	308010	8	56.0	86.3	92.0	-6.2	1
18	308012	8	68.0	67.7	88.0	-23.0	1
19	308023	8	92.0	55.6	76.0	-26.8	0
20	308024	8	92.0	85.5	92.0	-7.1	0
21	308029	8	72.0	64.4	74.0	-13.0	1
22	308031	8	56.0	68.5	94.0	-27.1	0
23	308034	8	52.0	62.9	92.0	-31.6	1
24	308036	8	72.0	83.9	96.0	-12.6	0
25	308037	8	96.0	52.4	66.0	-20.6	0
26	308038	8	36.0	62.9	78.0	-19.4	1
27	409004	9	88.0	59.7	80.0	-25.4	0
28	409006	9	96.0	64.5	78.0	-17.3	0
29	409009	9	88.0	66.9	84.0	-20.3	0
30	409013	9	64.0	57.3	64.0	-10.5	1
31	409028	9	64.0	46.8	58.0	-19.4	1
32	409029	9	48.0	50.0	58.0	-13.8	1
33	410001	10	48.0	74.2	92.0	-19.4	0
34	410002	10	48.0	58.9	80.0	-26.4	0
35	410004	10	52.0	77.4	90.0	-14.0	0

Notes: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively.
Deviation Index was measured by 100 x (Expected scores - Postscores) / Postscores).
Students with lower self-efficacy were identified when the deviation index was negative but above 5% in magnitude.

Table A14.1 (contd..2): Achievers Having Lower Self-Efficacy

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Deviation Index %	Sex (Boys=0, Girls=1)
36	410005	10	40.0	50.0	64.0	-21.9	0
37	410006	10	60.0	69.4	94.0	-26.2	0
38	410009	10	68.0	66.1	98.0	-32.5	0
39	410011	10	76.0	88.7	98.0	-9.5	0
40	410012	10	36.0	46.8	80.0	-41.5	0
41	410015	10	44.0	76.2	94.0	-19.0	0
42	410016	10	64.0	59.7	92.0	-35.1	0
43	410017	10	52.0	76.7	92.0	-16.6	0
44	410019	10	60.0	70.2	96.0	-26.9	0
45	410020	10	40.0	86.3	92.0	-6.2	1
46	410021	10	32.0	54.1	90.0	-39.9	1
47	410024	10	84.0	77.4	96.0	-19.4	1
48	410025	10	52.0	67.7	92.0	-26.4	0
49	410026	10	48.0	62.1	84.0	-26.1	0
50	410027	10	28.0	46.8	84.0	-44.3	0
51	410029	10	68.0	69.4	94.0	-26.2	0
52	410030	10	44.0	61.5	88.0	-30.2	0
53	410032	10	60.0	58.9	90.0	-34.6	1
54	410033	10	68.0	67.7	96.0	-29.4	0
55	410035	10	44.0	32.7	46.0	-28.8	0
56	410036	10	84.0	86.3	98.0	-11.9	1
57	410037	10	48.0	51.6	56.0	-7.8	0
58	410038	10	48.0	40.3	78.0	-48.3	0
59	410039	10	44.0	58.1	84.0	-30.9	0
60	410040	10	76.0	90.3	98.0	-7.8	0
61	410041	10	52.0	49.4	88.0	-43.8	0
62	513016	13	24.0	54.0	61.0	-11.4	0
63	614008	14	68.0	56.5	74.0	-23.7	0
64	614015	14	88.0	66.1	78.0	-15.2	0
65	614021	14	96.0	64.5	76.0	-15.1	0
66	614024	14	56.0	47.6	74.0	-35.7	0
67	614025	14	44.0	49.8	78.0	-36.2	0
68	615005	15	96.0	63.7	88.0	-27.6	0
69	615007	15	96.0	74.1	88.0	-15.8	0

Notes: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively.
Deviation Index was measured by 100 x (Expected scores - Postscores) / Postscores).
Students with lower self-efficacy were identified when the deviation index was negative but above 5% in magnitude.

Table A14.1 (contd..3): Achievers Having Lower Self-Efficacy

Sr.No	Student ID	Class ID	Prior Knowledge %	Self-Efficacy %	Achievement %	Deviation Index %	Sex (Boys=0, Girls=1)
70	615011	15	92.0	45.2	60.0	-24.7	0
71	615020	15	96.0	80.6	88.0	-8.4	0
72	615022	15	96.0	78.2	84.0	-6.9	0
73	615023	15	84.0	80.6	86.0	-6.2	0
74	615025	15	92.0	78.2	86.0	-9.0	0
75	615028	15	72.0	84.7	92.0	-8.0	0
76	615029	15	84.0	84.2	90.0	-6.4	0
77	615031	15	100.0	84.7	92.0	-8.0	0
78	616003	16	96.0	39.5	48.0	-17.7	1
79	616035	16	92.0	51.6	60.0	-14.0	1
80	616036	16	100.0	25.8	68.0	-62.0	1

Notes: Prior knowledge, self-efficacy and achievement were measured by prescores, expected scores and achievement postscores, respectively.
Deviation Index was derived by 100 x (Expected scores - Postscores) / Postscores).
Students with lower self-efficacy were identified when the deviation index was negative but above 5% in magnitude.

Appendix 15

DERIVATION OF MEDIATED EFFECTS

Table A15.1: All Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=464)

Model	WITHIN Class (Level 1) Explanatory Variables		Criterion Variable: Students' Achievement (%) ML Estimates of Regression Coefficients					Mediator Variable: SEI ML Estimates of Regression Coefficients				
			b	se	β	se	*p***	*d*	se	β (*d*)	se	*p*
SEM0a	Influence of Classroom Learning Environment (CLE)											
	Direct effect	CLE	6.767	2.028	0.218	0.066	0.000					
SEM0b	Mediation through Self-Efficacy (SEI):											
		CLE	2.978	1.750	0.096	0.059	0.001	8.595	2.577	0.270	0.076	0.000
	Mediator	SEI	0.435	0.079	0.448	0.057	0.000					
	Mediated effect*		3.739	1.311	0.121	0.037	0.000					
		t-values	2.853		3.237							

Note:* Using the product method, the mediated effect of CLE (unstandardised) was obtained as: 3.739= 0.435 x 8.595 and the corresponding standardised coefficient (β) was 0.121 = 0.448 x 0.270

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.1 (contd..2): All Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=464)

Model	WITHIN Class (Level 1) Explanatory Variables		Criterion Variable: Students' Achievement (%) (ACH)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	se	β	se	*p***	*d*	se	β (*d*)	se	*p*
SEM1a	Influence of Prior Knowledge of Type P1 (PK1)											
	Direct effect	PK1	0.199	0.044	0.383	0.067	0.000					
SEM1b	Mediation through Self-Efficacy (SEI):											
		PK1	0.120	0.034	0.231	0.057		0.182	0.037	0.345	0.066	0.000
	Mediator	SEI	0.409	0.068	0.415	0.043						
	Mediated effect*		0.074	0.020	0.143	0.031	0.000					
		t-values	3.808		4.596							

Note: * Using the product method, the mediated effect of PK1 (unstandardised) was obtained as: 0.074= 0.409 x 0.182, and the corresponding standardised coefficient (β) was 0.143 = 0.415 x 0.345.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.1 (contd..3): All Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=464)

Model	WITHIN Class (Level 1) Explanatory Variables		Criterion Variable: Students' Achievement (%) (ACH)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	se	β	se	*p***	*d*	se	β (*d*)	se	*p*
SEM2a	Influence of Prior Knowledge of Type P2 (PK2)											
	Direct effect	PK2	0.168	0.038	0.380	0.073	0.000					
SEM2b	Mediation through Self-Efficacy (SEI):											
		PK2	0.114	0.034	0.257	0.072	0.000	0.128	0.026	0.287	0.055	0.000
	Mediator	SEI	0.410	0.068	0.415	0.046	0.000					
	Mediated effect*		0.052	0.014	0.119	0.026	0.000					
		t-values	3.813		4.517							

Notes: * Using the product method, the mediated effect of PK2 (unstandardised) was obtained as 0.052 = 0.410 x 0.128, and the corresponding standardised coefficient (β) was 0.119 = 0.415 x 0.287

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx

Table A15.1 (contd..4): All Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=464)

Model	WITHIN Class (Level 1) Explanatory Variables		Criterion Variable: Students' Achievement (%) ML Estimates of Regression Coefficients					Mediator Variable: SEI ML Estimates of Regression Coefficients				
			b	se	β	se	*p*	*d*	se	β(*d*)	se	*p*
SEM3a	Influence of Prior Knowledge of All Types (PK)											
	Direct effect	PK	0.376	0.066	0.540	0.067	0.000					
SEM3b	Mediation through Self-Efficacy (SEI):											
		PK	0.257	0.053	0.371	0.063	0.000	0.305	0.055	0.444	0.069	0.000
	Mediator	SEI	0.374	0.063	0.370	0.041	0.000					
	Mediated effect*		0.114	0.028	0.164	0.031	0.000					
		t-values	4.052		5.239							

Notes: * Using the product method, the mediated effect of PK (unstandardised) was obtained as: 0.114 = 0.374 x 0.305 and the corresponding standardised coefficient (β) was 0.164 = 0.370 x 0.444.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Appendix 15 (contd..5)

Table A15.2: High Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=303)

Model	WITHIN Class (Lev Explanatory Variables		Criterion Variable: Students' Achievement (%) ML Estimates of Regression Coefficients					Mediator Variable: SEI ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM0a	Influence of Classroom Learning Environment (CLE)											
	Direct effect	CLE	8.475	2.763	0.316	0.083	0.000					
SEM0b	Mediation through Self-Efficacy (SEI)											
		CLE	5.635	1.925	0.210	0.059	0.000	9.402	3.244	0.304	0.099	0.002
	Mediator	SEI	0.297	0.045	0.342	0.036	0.000					
	Mediated effect*		2.792	1.052	0.104	0.036	0.016					
		t-values	2.654		2.922							

Note:* Using the product method, the mediated effect of CLE (unstandardised) was obtained as: 2.792 = 0.297 x 9.402 and the corresponding standardised coefficient (β) was 0.104 = 0.342 x 0.304.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.2 (contd..2): High Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=303)

Model	WITHIN Class (Level 1) Explanatory Variables		Criterion Variable: Students' Achievement (%) ML Estimates of Regression Coefficients					Mediator Variable: SEI ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM1a	Influence of Prior Knowledge of Type P1 (PK1)											
	Direct effect	PK1	0.198	0.061	0.353	0.091	0.000					
SEM1b	Mediation through Self-Efficacy (SEI)											
		PK1	0.154	0.053	0.276	0.082	0.001	0.131	0.053	0.207	0.084	0.014
	Mediator	SEI	0.304	0.042	0.345	0.028	0.000					
	Mediated effect		0.040	0.017	0.071	0.030						
		t-values	2.339		2.416							

Note:* Using the product method, the mediated effect of PK1 (unstandardised) was obtained as: 0.040 = 0.304 x 0.131 and the corresponding standardised coefficient (β) was 0.071 = 0.345 x 0.207.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.2 (contd..3): High Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=303)

Model	WITHIN Class Explanatory Variables		Criterion Variable: Students' Achievement (%)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients at Level 1					ML Estimates of Regression Coefficients at Level 1				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM2a	Influence of Prior Knowledge of Type P2 (PK2)											
	Direct effect	PK2	0.161	0.048	0.350	0.090	0.000					
SEM2b	Mediation through Self-Efficacy (SEI):											
		PK2	0.127	0.046	0.278	0.089	0.002	0.106	0.025	0.203	0.048	0.000
	Mediator	SEI	0.307	0.048	0.350	0.038	0.000					
	Mediated effect		0.033	0.009	0.071	0.018	0.000					
		t-value	3.534		3.843							

Note:* Using the product method, the mediated effect of PK2 (unstandardised) was obtained as: 0.033 = 0.307 x 0.106 and the corresponding standardised coefficient (β) was 0.071 = 0.350 x 0.203.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.2 (contd..4): High Ability Students: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=303)

Model	WITHIN Class at Level 1 Explanatory Variables		Criterion Variable: Students' Achievement (%)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM3a	Influence of Prior Knowledge of All Types (PK)											
	Direct effect	PK	0.422	0.095	0.441	0.083	0.000					
SEM3b	Mediation through Self-Efficacy (SEI**):**											
		PK	0.330	0.088	0.345	0.081	0.000	0.318	0.075	0.295	0.070	0.000
	Mediator	SEI	0.278	0.046	0.314	0.039	0.000					
	Mediated effect*		0.088	0.025	0.093	0.025	0.000					
		t-values	3.471		3.734							

Note:* Using the product method, the mediated effect of PK (unstandardised) was obtained as: 0.088 = 0.278 x 0.318 and the corresponding standardised coefficient (β) was 0.093 = 0.314 x 0.295.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx

Appendix 15 (contd..9)

Table A15.3: Low- and Medium-Ability Students Combined: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=161)

Model	WITHIN Class at Level 1 Explanatory Variables		Criterion Variable: Students' Achievement (%)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM0a	Influence of Classroom Learning Environment (CLE) (2nd Order CFA)											
	Direct effect	CLE	2.468	2.279	0.065	0.063	0.301					
SEM0b	Mediation through Self-Efficacy (SEI):											
		CLE	-0.076	2.574	-0.002	0.068	0.975	3.876	2.463	0.119	0.075	0.114
	Mediator	SEI	0.644	0.116	0.553	0.075	0.000					
	Mediated effect		2.496	1.649	0.066	0.042	0.123					
		t-values	1.514		1.551							

Note:* Using the product method, the mediated effect of PK (unstandardised) was obtained as: 2.496 = 0.644 x 3.876 and the corresponding standardised coefficient (β) was 0.066 = 0.553 x 0.119.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx

Table A15.3 (contd..2): Low- and Medium-Ability Students Combined: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=161)

Model	WITHIN Class at Level 1 Explanatory Variables		Criterion Variable: Students' Achievement (%) ML Estimates of Regression Coefficients					Mediator Variable: SEI ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM1a	Influence of Prior Knowledge of Type P1 (PK1)											
	Direct effect	PK1	0.163	0.060	0.156	0.059	0.009					
SEM1b	Mediation through Self-Efficacy (SEI):											
		PK1	0.031	0.060	0.029	0.060	0.625	0.202	0.076	0.223	0.084	0.008
	Mediator	SEI	0.630	0.115	0.541	0.076	0.000					
	Mediated effect		0.127	0.053	0.121	0.049	0.014					
		t-values	2.391		2.487							

Note:* Using the product method, the mediated effect of PK (unstandardised) was obtained as: 0.127= 0.630 x 0.202 and the corresponding standardised coefficient (β) was 0.121 = 0.541 x 0.223.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx

Appendix 15 (contd..11)

Table A15.3 (contd..3): Low- and Medium-Ability Students Combined: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=161)

Model	WITHIN Class at Level 1 Explanatory Variables		Criterion Variable: Students' Achievement (%)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM2a	Influence of Prior Knowledge of Type P2 (PK2)											
	Direct effect	PK2	0.100	0.167	0.039							
SEM2b	Mediation through Self-Efficacy (SEI):											
		PK2	0.077	0.042	0.129	0.072	0.073	0.038	0.063	0.074	0.124	0.551
	Mediator	SEI	0.636	0.106	0.544	0.067	0.000					
	Mediated effect		0.024	0.040	0.040	0.068	0.553					
		t-values	0.600		0.595							

Note: * Using the product method, the mediated effect of PK (unstandardised) was obtained as: 0.024 = 0.636 x 0..038 and the corresponding standardised coefficient (β) was 0.040= 0.544 x 0.074.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx.

Table A15.3 (contd..4): Low- and Medium-Ability Students Combined: ML Estimates of Direct and Mediated Effects of Prior Knowledge and Classroom Learning Environment on Students' Achievement with Self-Efficacy as Mediator from Mediation Model (N=161)

Model	WITHIN Class at Level 1 Explanatory Variables		Criterion Variable: Students' Achievement (%)					Mediator Variable: SEI				
			ML Estimates of Regression Coefficients					ML Estimates of Regression Coefficients				
			b	s.e	β	s.e	*p*	*d*	s.e	β (*d*)	s.e	*p*
SEM3a	Influence of Prior Knowledge of All Types (PK)											
	Direct effect	PK	0.503		0.353		0.000					
SEM3b	Mediation through Self-Efficacy (SEI)											
		PK	0.299	0.095	0.210	0.076	0.006	0.318	0.134	0.262	0.105	0.012
	Mediator	SEI	0.587	0.100	0.500	0.065	0.000					
	Mediated effect		0.187	0.085	0.131	0.055	0.019					
		t-values	2.200		2.373							

Note: * Using the product method, the mediated effect of PK (unstandardised) was obtained as: 0.187 = 0.587 x 0..318 and the corresponding standardised coefficient (β) was 0.131= 0.500 x 0.262.

** The standard errors of mediated effect were obtained by using the formula recommended by MacKinnon, Fairchild and Fritz (2007, p. 598) and *p*-values from http://www.socscistatistics.com/pvalues/tdistribution.aspx