SCHOOL OF EDUCATION

Investigating relationships between girls' confidence in their ability to do and learn mathematics, their perceptions of their learning environment and achievement.

Sharon Elizabeth Quinn

0000-0003-2426-5399

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

March 2021

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. 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.

The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number SMEC–39–14.

Signature:

Date: March 29 2021

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the many people who have made it possible for me to complete this thesis. I am most grateful for the help, advice and support they have given me during this incredible journey.

Thank you to my family and friends: my husband for his love and support; my children and grandchildren, from whom I have learnt so much about how we learn; my parents for instilling in me a love of learning; and my sister and my golfing buddies, who have provided encouragement and kept me on task.

Thank you to the staff and students at the schools who supported my work through their participation in the study: the staff who distributed and collected consent forms, and supervised students answering the questionnaires; the students who answered the questionnaires; and the school administrators who allowed this disruption to normal routines.

I am grateful to the Adelaide-based Curtin University group: Professor Barry Fraser, Professor Jill Aldridge and my fellow students. Our twice-yearly meetings provided opportunities for us to share each other’s triumphs and struggles. I always came away from these meetings with renewed vigour and purpose.

I would like to thank the members of my thesis committee, Dr Martin Cooper, Dr Jill Aldridge and Dr David Henderson, for their support. In particular, I thank my supervisor Professor Jill Aldridge, without whom this work would not have been completed. Her encouragement and gentle prodding together with her calm and positive attitude provided much needed inspiration when the task seemed overwhelming. Her expert guidance and constructive comments have enabled me to reach the end of the journey. Thank you.

I am grateful to Dr Kerrie Le Lievre who copy-edited my thesis according to the Australian Standards of Editing Practice.

This research was supported by an Australian Government Research Training Program (RTP) scholarship.

ABSTRACT

In light of the frequent calls to increase students', particularly girls', participation in STEM subjects, this thesis focusses on exploring associations between female students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The study also examines the existence of relationships between confidence and achievement, as well as differences between students' levels of confidence and their perceptions of the classroom leaning environment based on membership of particular groups.

Quantitative research methods were used. Data were collected using two questionnaires which were modified for use in this study: the Attitudes Towards Mathematics Inventory (m-ATMI) which was used to assess students' confidence in their ability to do and learn mathematics, and the What Is Happening In this Class? (WIHIC) questionnaire, which was used to assess students' perceptions of the classroom learning environment. The sample comprised 335 female students in 25 classes, at two South Australian girls' schools. Students at School 1 completed the questionnaires on three separate occasions, each time at the end of the completion of a unit of work. Students at School 2 completed the questionnaires once. Student achievement grades based on end of unit tests were provided by the teachers.

The structures of the questionnaires were examined separately using principal axis factoring with oblique rotation. The criterion for retention of an item was that it must have a factor loading of at least .40 on its own scale and less than .40 on the other scales. This resulted in the retention of all items of the m-ATMI and the removal of the Equity scale and seven other items from the WIHIC. Cronbach's alpha coefficient was calculated to measure the internal consistency reliability of each scale. The component correlation matrices were examined to determine the discriminant validity of each questionnaire. All of these analyses supported the satisfactory factorial and discriminant validity and the internal consistency reliability of both the m-ATMI and the WIHIC for use with middle-school girls in South Australia .

Simple and multiple regressions were used to investigate the relationships between the variables confidence, classroom learning environment and achievement. Confidence

was found to be positively and significantly related to all scales of the WIHIC, with Clarity of Assessment Criteria and Involvement having the largest correlations. Confidence was also strongly correlated with achievement.

Multivariate analysis of variance (MANOVA) was used in order to investigate the statistical significance of differences in confidence and perception of the classroom learning environment based on the independent variables year level (Years 7 to 10), ability group (extension vs mainstream) and strand of mathematics being studied (Number, Algebra, Measurement, Trigonometry and Statistics). The results showed that although confidence decreased as students progressed through school, the differences were not statistically significant. There were, however, statistically significant differences found between year levels for the Teacher Support, Cooperation and Involvement scales of the WIHIC. Confidence levels were statistically significantly higher for students in Extension Mathematics classes. These students also had better perceptions of the classroom learning environment, with statistically significant differences being present for the Clarity of Assessment Criteria, Teacher Support and Involvement scales of the WIHIC. There were no statistically significant differences found for confidence levels based on the strand of mathematics being studied.

The study is significant in that the questionnaires validated and used in this study can be used by other researchers and teachers to assess middle-school girls' confidence in their ability to do and learn mathematics and their perceptions of the mathematics classroom learning environment. This study is different from other research into mathematics education because the focus is on affective rather than cognitive aspects of learning. As such, it adds to the growing body of knowledge with regard to relationships between elements of classroom learning and self-concept-like constructs such as confidence. The results reported in this thesis have implications for mathematics teachers in their quest to increase girls' confidence in their ability to do and learn mathematics. Improving girls' confidence in their ability to do and learn mathematics is important because lack of confidence is often cited as a reason for their lack of participation in higher-level mathematics courses at school and their subsequent involvement in STEM careers. The study supports the view that one factor in improving girls' confidence in their ability to do and learn mathematics is related to the classroom learning environment.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ACRONYMS

AAMT	Australian Association of Mathematics Teachers
ACARA	Australian Curriculum, Assessment and Reporting Authority
ANOVA	Analysis of Variance
ATMI	Attitudes Towards Mathematics Inventory
CES	Classroom Environment Scale
CFA	Confirmatory Factor Analysis
CLEI	Chemistry Learning Environment Inventory
CLES	Constructivist Learning Environment Scale
COLES	Constructivist-Oriented Learning Survey
CUCEI	College and University Classroom Learning Environment Inventory
FSMAS	Fennema-Sherman Mathematics Attitudes Scale
HSD	Honestly Significant Difference
IB-MYP	International Baccalaureate Middle Years Programme
ICEQ	Individualized Classroom Environment Scale
ICSEA	Index of Community Socio-Educational Advantage
KMO	Kaiser-Meyer-Olkin
LEI	Learning Environments Inventory
MANOVA	Multivariate Analysis of Variance
m-ATMI	modified Attitudes Towards Mathematics Inventory
MATS	Mathematics and Technology Scale
MCI	My Class Inventory
OECD	The Organization for Economic Co-operation and Development
PALS	Patterns of Adaptive Learning Scales
PISA	Programme for International Student Assessment
QSL	Quality of School Life
QTI	Questionnaire on Teacher Interaction
SACSA	South Australian Curriculum Standards and Accountability

SALES	Student Adaptive Learning Engagement in Science
SDQ	Self-Description Questionnaire
SES	Socio-Economic Status
SLEI	Science Laboratory Environment Scale
SLEQ	Scholl Level Environment Questionnaire
SPAQ	Students' Perceptions of Assessment Questionnaire
STEM	Science, Technology, Engineering and Mathematics
TIMSS	Trends in Mathematics and Science Study
TOSRA	Test of Science Related Attitudes
TROFLEI	Technology-Rich Outcomes-Focused Learning Environment Inventory
UAE	United Arab Emirates
WIHIC	What Is Happening In this Class?

Chapter 1

INTRODUCTION

1.1 Background to the Study

For some years, in Australia, government, academic and business leaders have been calling for greater participation in Science, Technology, Engineering and Mathematics (STEM) subjects by all school students (Campanini, 2020; Office of the Chief Scientist, 2013, 2014; Validakis, 2015). These calls to increase student participation are usually accompanied by statistics showing the low numbers of girls, when compared with boys, studying higher-level mathematics courses at school and subsequently pursuing STEM careers (Campanini, 2020; Kennedy, Lyons, & Quinn, 2014; Li & Koch, 2017).

During the last fifty years there has been a considerable amount of research into girls and mathematics. It has focussed on attitudes, participation rates and achievement, and has consistently identified a gender gap in favour of boys in all areas. The consequence of these gender gaps is that females are underrepresented in the STEM workforce, making up "roughly 50% of the general population but only 25% of the overall STEM workforce" (Ellis, Fosdick, & Rasmussen, 2016, p. 2). As mathematics is considered a prerequisite subject for study in all STEM areas and

> school learning in mathematics and science subjects during the middle school years often contributes to student selection of STEM disciplines in upper secondary, shaping their pathways into tertiary study (Roberts, 2014, p. 18)

it is important to find ways to increase female participation in school mathematics courses.

Research into attitudes towards mathematics has consistently found that girls are less confident in their ability to do and learn mathematics than boys. This lack of confidence is the most common reason proposed for girls' poor rates of participation in advanced mathematics courses (Fennema & Sherman, 1978; Hyde, Fennema, Ryan, Frost, &

Hopp, 1990; Lin & Huang, 2014; Nurmi, Hannula, Maijala, & Pehkonen, 2003; Thomson et al., 2012; Watt, 2007; Watt et al., 2017). The reasons suggested to explain their lack of confidence are numerous, and include factors such as peers, classroom learning environments, family, and society stereotypes. The only one of these factors which the classroom teacher has any direct influence over is the classroom learning environment. Elizabeth Fennema therefore argues (2000, p. 2) that in order to achieve equity of outcomes for girls and boys, it is necessary to modify the classroom environment to suit all students, not just particular groups.

Classroom learning environment research has found links between the classroom learning environment and both cognitive and affective educational outcomes. It has consistently shown that students achieve better outcomes (both cognitive and affective) in their preferred classroom learning environments (Fraser, 2012b, 2014). Extensive research has also shown that "creating positive classroom environments is likely to pay off in terms of improving student outcomes" (Fraser, 2019, p. 2). There is a large body of evidence supporting the existence of relationships between classroom learning environments and affective outcomes such as attitude, confidence and academic efficacy (Aldridge, Afari, & Fraser, 2012; Aldridge & Fraser, 2008; Dorman, 2001; Dorman & Adams, 2004; Dorman, Fisher, & Waldrip, 2006; Yang, 2015).

Another factor that teachers link with confidence is achievement. Achievement implies the accomplishment of something; in schools, it is associated with "articulated learning goals" (Guskey, 2013). While learning goals have cognitive, affective and psychomotor elements, the basis of most academic achievement is cognitive. It generally relates to a specific subject (e.g., mathematics) and describes the concepts and skills that teachers consider important for students to acquire. While acknowledging that learning is not confined to classrooms, schoolteachers focus on the learning that takes place within their classrooms. It seems logical to assume that there is a relationship between confidence and achievement, and yet research has reported inconsistent findings (Hattie, 1992; Huang, 2011; Ma & Kishor, 1997b), indicating that the relationship is complicated.

From my earliest experiences with teaching mathematics to girls, lack of confidence was presented as a problem with respect to their success at and participation in mathematics. I was interested to see whether there were any statistically significant relationships between girls' confidence in their ability to do and learn mathematics and their perceptions of their classroom learning environment, and what role achievement might play in determining their confidence in their ability to do and learn mathematics. Through the use of questionnaires, it was possible to elicit information from students about their confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The main focus of this study was on determining whether aspects of the classroom learning environment – something that teachers can influence – have any relationship to female students' confidence in their ability to do and learn mathematics.

This chapter introduces the problem to be address by this study: Can teachers change girls' confidence in their ability to do and learn mathematics by making changes to the classroom learning environment? It begins by describing the context of the study (Section 1.2) , its theoretical underpinnings (Section 1.3) and its conceptual framework (Section 1.4). It then introduces the main objectives of the study (Section 1.5) and provides an overview of the significance of the research (Section 1.6). An overview of the thesis is provided in Section 1.7.

1.2 Context of the Study

This section describes the context in which this study was conducted. Section 1.2.1 provides general information about South Australia and education in South Australia, while Section 1.2.2 provides information about the two schools from which the sample in the study was drawn.

1.2.1 Education in South Australia

To provide context for the study, this section briefly describes South Australia and its education systems. South Australia is one of the eight states and territories of Australia and is situated in the southern central region. The population of South Australia is approximately 1.7 million, with 77% living in the capital city of Adelaide. South Australia was founded as a British colony in 1836; consequently, the majority of South Australians are of European ancestry. Aboriginal and Torres Strait Islander people comprise approximately 2% of the population. During the late twentieth century there was a significant increase in immigration from Asian and other non-European countries, creating a multicultural society.

The responsibility for providing education to Australian children rests with the states and territories. They meet their commitment to universal access to education through the provision of public education. Two systems operate within South Australia: the public sector (which is mainly funded by the state government) and the private sector (which is largely funded by the Commonwealth government). Schools in the private sector are primarily church-based, and rely on parent contributions and government funding (Harrington, 2011). This sector has two broad categories of schools: Catholic and non-Catholic. In South Australia, Catholic schools are either diocesan or congregational; the state Catholic Education Office is responsible for the governance of diocesan schools, while the responsibility for governance of congregational schools lies with the school board or council, which is appointed by the congregation. The schools used in this research were both congregational.

The curriculum in most Australian schools is based on a national curriculum developed by the Australian Curriculum, Assessment, and Reporting Authority (ACARA), which is responsible for the administration of national assessments and the associated reporting on schooling in Australia. Schools are responsible for the delivery of the curriculum, assessment of students and reporting of student outcomes to parents. The states are responsible for how the national curriculum content and achievement reporting are integrated into school programmes. In South Australia, the Education Standards Board is responsible for ensuring that all schools, public and private, effectively deliver the Australian Curriculum or an approved equivalent. The Board considers each school's suitability for registration based on three broad areas: school governance, studying learning and assessment, and student safety, health and welfare. Registration is reviewed every five years. Other curriculum frameworks, such as the International Baccalaureate Middle Years Programme (IB-MYP), can be approved for use in schools providing that they meet the requirements of the national curriculum.

The Australian Curriculum: Mathematics is designed to provide students with "essential mathematical skills and knowledge" (ACARA, 2010, p. 4). It focusses on developing students' mathematical understanding in order to provide them with the ability to use mathematics in a variety of situations and see how it is related to other disciplines. The curriculum is organised around three content and four proficiency strands: the content strands (Number and Algebra, Measurement and Geometry, and Statistics and Probability) describe what is to be taught and learnt, and the proficiency strands (understanding, fluency, problem solving, and reasoning) describe how the content is explored or developed (ACARA, 2010, p. 6). Achievement standards are prescribed for each year level.

1.2.2 The Sites Used

The study reported in this thesis was conducted at two Catholic girls' schools in the eastern suburbs of Adelaide, South Australia. Although the schools share the same postcode, during the period in which the study was carried out they had different Socio-Economic Status (SES) and Index of Community Socio-Educational Advantage (ICSEA) scores. SES scores are used to measure a school's ability to support itself and determine the level of government funding provided to the majority of private schools. They are calculated using census data about families' incomes, education and occupations. ICSEA scores are used to measure a school's level of educational advantage, and are calculated using parent occupations and levels of education completed. They are intended to provide parents with the means to compare a school's performance in literacy and numeracy with that of other schools that have similar backgrounds. School 1 scored higher on both indices, with an SES score of 113 and an ICSEA score of 1131, which placed it in the ninetieth percentile of schools. School 2's scores were SES 102 and ICSEA 1039, placing it in the sixty sixth percentile. Another difference at the time of the study was that School 2 had a much higher percentage of students from non-English speaking backgrounds than School 1: 98% compared to 10%. The differences in family education levels, occupations, income and language backgrounds provided greater diversity in the sample.

School 1, with an ICSEA score of 1131, was a Reception to Year 12 boarding school with approximately 800 students on 2 campuses. At the time of the study the IB-MYP was the framework used to determine curriculum and methodology. The IB-MYP is designed to "encourage the use of a variety of teaching and learning methodologies fostering a climate in which students discover how they learn best in different situations" (International Baccalaureate Organization, 2013). For mathematics, the school provided two courses, Mathematics and Extended Mathematics, which both identify five strands of mathematics: Number, Algebra, Geometry and Trigonometry, Statistics and Probability, and Discrete Mathematics. In this school, the Extended Mathematics programme was offered to between 15% and 20% of students in each year level. In Years 6 to 9, selected students were withdrawn from their normal mathematics class once a week for extension activities based on the Extended Mathematics curriculum. In Year 10 there was a separate class which used the IB-MYP extended mathematics curriculum and was intended to prepare students for the study of two units of mathematics in Year 12. All other students were randomly assigned to classes.

School 2, with an ICSEA score of 1039, was a Year 8 to 12 day school with approximately 350 students. This school used the Australian Curriculum to determine class content. The core subjects of the Australian Curriculum, such as mathematics, give students experiences in a variety of disciplines and allow for the development of required knowledge and skills. For mathematics, students were provided with essential mathematical skills, knowledge and understanding in three strands: Number and Algebra, Measurement and Geometry, and Statistics and Probability. The Australian Curriculum: Mathematics is designed to develop students' numeracy capabilities and to provide them with the basics on which future mathematics will rely. It aims to place student learning and independence at its centre. The content descriptions specify what teachers are expected to teach. They provide a well-researched scope and sequence of teaching, within which teachers determine how best to cater for individual students' learning needs and interests. An achievement standard describes the quality of learning (i.e., depth of understanding, extent of knowledge and sophistication of skill) typically expected of students as they progress through schooling (ACARA, 2010). In this school, students in Years 8 to 10 were randomly assigned to mathematics classes, with a small group of students at Year 10 provided with a less rigorous course.

In both schools, student achievement was measured using a combination of assessment tasks, including common tests, end-of-semester tests and research projects. The greatest weight was given to the common test, so this was chosen as the measure to be used in this study. It was also the least subjective form of assessment used in these schools, and so the most suitable for making comparisons.

1.3 Theoretical Underpinnings

This section identifies the theoretical frameworks which underpin this research. The study draws upon three areas of study in examining the relationships between girls' confidence in their ability to do and learn mathematics, their perceptions of the classroom learning environment and achievement. These are self-perceptions, the gender gap and learning environments.

This study draws extensively on the field of self-perceptions in examining the relationship between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The first modern theory of self was developed by James (1890) when he described self as the sum as all aspects of a human being, implying that self is multi-faceted and hierarchical. The multi-faceted and hierarchical nature of self has resulted in a plethora of terms such as self-concept, self-efficacy, academic efficacy and confidence which are often used interchangeably. The term of interest to this study is confidence which has been described as an attitude (Fennema & Sherman, 1976, 1977; Reyes, 1984; Reyes & Stanic, 1988; Tapia & Marsh, 2004), an affective variable (Cretchley, 2008), a belief (Benabou & Tirole, 2002; Druckman & Bjork, 1994) and a personality trait (Kleitman & Stankov, 2007; Stankov, 1999; Stankov, Kleitman, & Jackson, 2015). Self-perceptions such as confidence are important in the development of motivation and learning outcomes (Bandura, 1986; Jinks & Lorsbach, 2003; Zimmerman, 2000).

Lack of confidence has long been linked with girls' under-representation in STEM subjects and careers (Forgasz, Leder, & Kloosterman, 2004; Ganley & Lubienski, 2016; Lubienski, Robinson, Crane, & Ganley, 2013; Watt, 2006; Watt, Eccles, & Durik, 2006; Zeldin & Pajares, 2000). Lower participation rates by girls in STEM subjects at school and later STEM careers forms the basis for another area of research relevant to this study, known as the gender gap. Before the 1970's it was accepted that more males than females would study mathematics and that generally male achievement levels would be higher than those of females. Since then a huge amount of time and money has been invested in programmes aimed at redressing these imbalances. Sadly, while the achievement gap has narrowed, women and girls are still underrepresented in higher-order mathematics courses and subsequently STEM careers.

This study also draws heavily on the field of learning environments and its relationship to student outcomes such as confidence and academic efficacy. Past research in this area has focussed on measuring learning environments from the students' perspective, with Fraser (2001) indicating that students are well placed to do this because of the many hours they spend in the classroom. Learning environments research has examined associations between classroom learning environment and student outcomes, the effects of gender, year level and subject on the classroom learning environment, the transition from primary to secondary school and has been used to help teachers to improve their classroom learning environment. Consequently, it has provided researchers with valuable insights into the learning process (Chionh & Fraser, 2009; Fraser, Aldridge, & Adolphe, 2010; Ogbuehi & Fraser, 2007; Velayutham, Aldridge, & Fraser, 2011, 2012; Wolf & Fraser, 2008).

By recognising the important role that the environment plays in determining human behaviour early psychologists such as Lewin (1936) and Murray (1938) provided the basis for research into classroom learning environments. A range of questionnaires were developed to assist in this research beginning with the Learning Environment Inventory (LEI) (Walberg & Anderson, 1968) and the Classroom Environment Survey (CES) (Trickett & Moos, 1973). A leading proponent of learning environment research in Australia, Barry Fraser, has shifted the focus from teacher-centred classrooms to student-centred classrooms. He has been involved in the development of a number of questionnaires used in this research, including the Individualized Classroom Environment Survey (ICEQ) (Fraser & Butts, 1982; Rentoul & Fraser, 1979); the Constructivist Learning Environment Survey (CLES) (Taylor, Fraser, & Fisher, 1997) and the What is Happening in this Class (WIHIC) (Aldridge, Fraser, & Huang, 1999; Fraser, McRobbie, & Fisher, 1996). Numerous other questionnaires have been since been developed for use in a range of research programmes across the world.

More detail about these areas of research which form the theoretical underpinnings of this study can be found in Chapter 2.

1.4 Conceptual Framework

This study was conducted using quantitative research methods. Quantative research methods sit within a positivist paradigm which can be described as “an epistemological position that advocates the application of the methods of the natural sciences to the study of social reality and beyond” (Bryman, 2012, p. 28). The value of quantitative research is that it is possible to get answers to questions and to generalise findings. Questions such as *How many students study mathematics? Is the number of students taking mathematics courses reducing?* can be answered using descriptive statistics, while questions such as *What factors are related to changes in student participation in mathematics? Is there a relationship between achievement and confidence?* can be answered using inferential statistics. Using quantitative research methods, it is possible to collect data from a large number of people and to compare groups or relate variables using statistical analysis. The results provide a snapshot of a particular point in time, that can then be compared with past research. A range of research designs are available to achieve this, including experimental designs, correlational designs and survey designs (Creswell, 2012). Experimental designs are used to establish possible causes and effects between variables, correlational designs are used to establish relationships between variables, and survey designs are used to describe the attitudes, opinions, behaviours and characteristics of a population.

A correlational research design was chosen for this study, as the main focus was on examining relationships between variables such as students’ confidence in their ability to do and learn mathematics, their perceptions of the classroom learning environment and their achievement scores in mathematics (L. Cohen, Manion, & Morrison, 2013; Creswell, 2012; Mertens & Wilson, 2012; Punch, 1998). There are two main types of correlational design: the explanatory design and the prediction design. The explanatory design is interested in the degree to which variables are related, while prediction design is interested in identifying variables which will predict future behaviour. This study utilised the explanatory design, which is preferable when “the researcher is interested in the extent to which two variables (or more) co-vary” (Creswell, 2012, p. 340).

In quantitative research methods, data are collected using instruments such as questionnaires, standardised tests and check lists. The instruments are used to measure observe and document the data, which are then analysed with a view to generalising to larger populations. This study made use of two well-known instruments to collect data about students' perceptions of their confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment: the Attitudes Towards Mathematics Inventory (ATMI) and the What Is Happening in this Class? questionnaire (WIHIC). Questionnaires are useful because they are an easy way to gather data that enables researchers to describe students' perceptions through their own eyes. Data collection no longer has to rely on third party observers, but is in the hands of the participants. As Fraser (2001) points out, students are a good source of data about the classroom learning environment because they have experienced a variety of classrooms, and hence learning environments, over their approximately 20,000 hours of schooling. Measures of the classroom learning environment "involve judgements of psychological or social-psychological states of classes" (Fraser & Walberg, 1991, p. x) rather than being measures of achievement or ability.

1.5 Research Questions

This section outlines the intention of this study, which was to answer the question "What are the relationships between female students' confidence in their ability to do and learn mathematics, the classroom learning environment and achievement?" This was derived from the frequent assertion that girls are less confident about their ability to do and learn mathematics than boys, and that this lack of confidence is one of the main reasons for their reduced participation in higher-level mathematics courses (Thomson et al., 2012; Watt, 2007; Watt et al., 2017). In order to collect data, two well-known instruments were modified for use with the current sample; to have confidence in the results obtained, it was necessary to first determine the validity and reliability of these instruments for use with the study's sample of South Australian schoolgirls. This led to the study's first research question.

Research Question 1:

> Are the surveys to be used to collect data for the present study valid and reliable for use in girls' Catholic schools in metropolitan South Australia?

The study's second research question pertains to the relationships between the students' confidence in their ability to do and learn mathematics, their perceptions of the classroom learning environment and their mathematics achievement scores.

Research Question 2:

> Do relationships exist between students' confidence in their ability to do and learn mathematics and their
>
> a. perceptions of their classroom learning environment?
> b. achievement scores in mathematics?

The third and fourth research questions examine whether students in different year levels and different ability groups have different perceptions of their confidence in their ability to do and learn mathematics, and of the classroom learning environment.

Research Question 3:

> Do students in different year levels differ in terms of their
>
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their classroom learning environment?

Research Question 4:

> Do students in extension classes differ from other students in the same year level in terms of their
>
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their classroom learning environment?

The final research question examines the possibility of students having different levels of confidence for different strands of mathematics (e.g., Algebra, Statistics).

Research Question 5:

Do confidence levels differ for different strands of mathematics?

1.6 Significance of the Research

This section provides an overview of the significance of the research reported in this study. More detail is provided in Chapter 5.

Firstly, lack of confidence has been shown to be one reason for girls not taking higher-level mathematics courses. Therefore, it is hard to dispute the importance of this study if, as expected, it provides a means for classroom teachers to help redress the imbalance in girls' participation in higher-level mathematics courses and subsequently STEM careers.

Secondly, the research provides evidence to support the reliability and validity of modified versions of the instruments used in the study (one to assess students' attitudes towards mathematics and the other to assess students' perceptions of the learning environment) when used in middle-school mathematics classes in two South Australian girls' schools. Thus, it provides researchers and teachers with two validated questionnaires that can be used both in future research and to measure the success of interventions designed to improve girls' confidence in their ability to do and learn mathematics.

Thirdly, most mathematical education research has considered the cognitive rather than the affective aspects of learning, and studies on confidence in mathematical ability have focused on increasing confidence by examining methodology, course content, use of technology and gender differences. This study is significant because it explores associations between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. It is one of relatively few studies that have explored associations between students' confidence in their ability to do and learn mathematics, their perceptions of the classroom learning environment and achievement. By identifying the psychosocial elements of the classroom learning environment, the research findings have the potential to highlight for teachers the influence of the classroom learning environment on confidence, and provide evidence that time spent creating a positive classroom learning environment is not wasted, because it is likely to create more positive student attitudes, which in turn should enhance girls' confidence in their ability to do and learn mathematics.

Finally, this study is different from other studies in that it investigates the differences in girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment based on year levels and ability groupings. It is also, as far as I can determine, the only study to investigate differences in confidence due to the strand of mathematics being studied. The study also has the distinction of being one of very few that has used the term confidence rather than other self-belief terms, making it more relevant to teachers because it uses terminology with which they are familiar.

1.7 Thesis Overview

The research that formed the basis for this thesis is reported in five chapters. Chapter 1 provides background and context for the study. It also outlines the conceptual framework used in completing the research. The study's research questions are introduced and the significance of the research is outlined. Finally, the chapter provides an overview of the thesis.

Chapter 2 presents a review of the literature relevant to this study. It begins by considering the gender gap with regard to girls' participation rates and achievement levels in mathematics. The chapter then reviews literature relating to three self-beliefs: self-concept, confidence and self-efficacy. It examines their similarities and differences and the confusion that exists around their usage, concluding by giving the definition of confidence used in this study. The chapter goes on to review literature on measuring students' confidence in their ability to do and learn mathematics, including details of some of the instruments which have been used in the past. There is a detailed discussion of the instruments used in this study (ATMI and WIHIC). Following this there is a review of literature on the existence of relationships and differences between confidence and the variables of interest: achievement, year level, ability grouping and strand of mathematics. Finally the chapter reviews literature related to classroom learning environments, starting with the historical background to this field of study and including descriptions of a number of instruments commonly used in research. Detailed information is provided about the instrument chosen for this study. An overview of past research involving classroom learning environment instruments, and in particular associations between classroom learning environments and student outcomes and determinants of classroom learning environments, concludes the chapter.

Chapter 3 outlines the research methods and procedures used in this study. It describes the sample used for the research and provides details of the pilot study conducted to test the questionnaires. It also contains information about the instruments used to collect the data, and provides details of the changes made to each of the instruments to ensure their suitability for use in this study. The chapter includes an outline of the data analysis methods to be used, and details of the ethical considerations involved in the research.

Chapter 4 reports details of the data analysis and the results for each of the study's research questions. It begins by providing evidence for the reliability and validity of the instruments used. It then provides results for the relationships between variables (i.e., confidence, classroom learning environment scales and achievement) as per the research questions. Using the scales of the m-ATMI and the WIHIC as dependent variables, it explores the differences between students' perceptions of their confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment at different year levels, in different ability groups and for the different strands of mathematics being studied.

Chapter 5 summarises the thesis and discusses the results for each of the research objectives outlined above. It also provides details of the study's limitations, with recommendations for future research. There is more information given about the significance of the study, as well as some advice to teachers on how they can assist in improving girls' confidence in their ability to do and learn mathematics.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides a review of literature related to the present study. The literature review begins in Section 2.2, which examines literature relating to the gender gap in mathematics participation and highlights some historical aspects of the issue and why it is important to reduce the gap. Section 2.3 reviews literature relating to self-beliefs, in particular confidence, which is the focus of this study. The review focuses on two broad categories of self-belief: self-concept-like and self-efficacy-like beliefs as described by Marsh et al. (2019). Some instruments which have been developed to measure self-concept-like perceptions such as confidence are described, with emphasis on the ATMI, the instrument used in this study. Finally, this section also reviews literature relating to associations between confidence and a number of variables employed in this study: achievement, year level, ability grouping and strand of mathematics being studied. Section 2.4 then provides an overview of the history and development of research in the area of classroom learning environments. It describes several instruments which have been developed to measure different aspects of the classroom learning environment, including the WIHIC. Lastly, it reviews research related to confidence and classroom learning environments. Section 2.5 provides a summary of the chapter.

2.2 Girls' Participation and Achievement Levels in Mathematics

The focus of this study is the relationship between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The reasons for focussing on girls is that it is generally accepted that women are underrepresented in STEM careers, which has long-term economic implications for them (Akerman, 2016; Blickenstaff, 2005; Campanini, 2020; Catsambis, 2005; Smith, 2010; Stoet & Geary, 2018). The causes of this underrepresentation begin at school, when girls do not choose the higher-order mathematics courses that provide a pathway to STEM careers. Lack of confidence is frequently cited as a reason for girls not choosing these mathematics courses, or STEM careers (Forgasz et al.2004; Ganley & Lubienski, 2016; Lubienski et al., 2013; Watt, 2006; Watt et al., 2006; Zeldin & Pajares, 2000). This section outlines historical research on girls' participation in mathematics, and the gender gap first identified by Elizabeth Fennema in the 1970s. It concludes by showing that while the achievement gender gap may have diminished, the participation gender gap has not.

Interest in girls' achievement and participation in mathematics was kick-started in the 1970s when Elizabeth Fennema asked the question: "Sex differences in mathematics-learning: Why?" (Fennema, 1974). Based on the limited data available at the time, she concluded "that high-school boys achieve at higher levels than high-school girls and that the difference in achievement increases as adolescence progresses" (Fennema, 1974, p. 183). Over the next 25 years she and her colleagues considered numerous possibilities in an attempt to find reasons for the gender differences in both achievement and participation in mathematics courses (Fennema, 2000; Fennema & Carpenter, 1998; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Fennema & Sherman, 1976, 1977, 1978). While numerous answers have been proposed, such as differences in spatial ability (Nuttal, Casey, & Pezaris, 2005), cognitive style (Arnup, Murrihy, Roodenburg, & McLean, 2013), verbal ability (Wang, Eccles, & Kenny, 2013) and gender-role stereotypes in mathematics (Fennema & Sherman, 1977; Forgasz et al., 2004; Hyde et al., 1990), the most common reason put forward is differences in self-beliefs such as self-efficacy (Pajares, 2005) and confidence (Ellis et al., 2016; Watt, 2005, 2007; Watt et al., 2006).

There is conflicting evidence about the existence of an achievement gender gap. A meta-analysis of studies published between 1990 and 2007 by Lindberg, Hyde, Petersen, and Linn (2010) revealed that there were no gender differences in mean performance between girls and boys. In contrast, a study based on the analysis of 20,000 kindergarten to fifth-grade students in the US (Fryer & Levitt, 2010), found a difference in performance that favoured boys, and also showed that the difference increased with age. The Organization for Economic Co-operation and Development (OECD), using the 2012 Programme for International Student Assessment (PISA) results, has also reported that "boys outperform girls in mathematics" (OECD, 2015, p. 64) and that the gap is "wider among top-performing students" (OECD, 2015, p. 68).

For Australian students the Trends in Mathematics and Science Study (TIMSS) 2015 data revealed a gender difference in mathematics achievement at Year 4 level for the first time since 1995, though there was no difference at the Year 8 level (Thomson, Wernett, O'Grady, & Rodrigues, 2017). A clue to the reason for these results may lie in the results of a US study (Reardon, Kalogrides, Fahle, Podolsky, & Zárate, 2018, p. 11) which recently reported that boys achieve higher scores when "a larger proportion of the overall score is based on multiple-choice items", which is what the PISA and TIMSS tests use.

While the gender gap in terms of achievement may still exist, it is small by comparison with the gender gap in participation rates for girls in higher-level mathematics courses (Barrington & Evans, 2016; Forgasz, 2012; Kennedy et al., 2014; Li & Koch, 2017; Rennie, 2010; UNESCO, 2012). In Australia, while there has been an increase in the number of Year 12 students enrolling in mathematics, reported "participation in intermediate and advanced mathematics remains at historic lows. The gender gap also remains, with more boys enrolled in intermediate and, particularly advanced, mathematics than girls" (Li & Koch, 2017, p. 2). It is also reported by Li and Koch (2017) that girls are 25% less confident in their mathematical ability than boys.

Social cognitive researchers and others have proposed girls' and women's self-beliefs about their capabilities (low self-confidence) as a reason for their underrepresentation in school mathematics and STEM-related careers (Bandura, 1997; Eccles, 2011a; Ellis et al., 2016; Hackett & Betz, 1989; Hart, 1989; Pajares, 1996; Reyes & Stanic, 1988; Watt, 2007). Eccles stated that:

> With regard to the gender difference in the occupations linked to math and physical science in particular, females are less likely to enter these fields than males both because they have less confidence in their math and physical science abilities and because they place less subjective value on these fields than they place on other possible occupational niches. (Eccles, 2011a, p. 195)

These views are supported by the OECD (2015, p. 64), which reports that "one factor that may hold girls back is confidence in their own abilities in mathematics". The report further states that “studies show that the learning environment plays a significant role in fostering, or undermining, girls’ sense of self-confidence”. The study reported in this thesis therefore examines the relationships between girls’ confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment.

This section has provided an overview of research into the gender gap in mathematics, beginning with the work of Elizabeth Fennema and her colleagues in the 1970s. Over time it has become clear that while the gender gap in achievement still exists, it has narrowed, and what remains may in part be a product of the type of tests used to measure achievement. The gender gap in participation rates, however, is still evident. One of the many reasons proposed for the existence of this gap is that girls lack confidence in their ability to do and learn mathematics. Most of the early research described above focussed on identifying confidence as one of the causes of the gender gap, rather than looking for ways to increase girls’ confidence in their ability to do and learn mathematics. The study reported in this thesis fills this gap by identifying ways in which girls’ confidence in their ability to do and learn mathematics may be increased.

2.3 Confidence

The study reported in this thesis explores the relationships between girls’ perceptions of the classroom learning environment and their confidence in their ability to do and learn mathematics. Self-perceptions such as confidence play a critical role in the development of motivation and learning outcomes (Bandura, 1986; Jinks & Lorsbach, 2003; Zimmerman, 2000). As such, confidence is one of many aspects of self that are studied today; others include self-concept, self-efficacy and academic efficacy.

Humans have always had an interest in the notion of self, beginning with the Greek philosophers Socrates and Plato (Hattie, 1992). One of the first modern theories of self was developed in the 1890s by William James, who stated, “Our self-feeling in this world depends entirely on what we back ourselves to be and do” (James, 1890, p. 200). For him a person’s self was “the sum total of all that he can call his” (James, 1890, p. 188) including his body, emotions, feelings, possessions, family, friends and work, which are hierarchically ordered. This suggests that self is a multi-faceted, hierarchical construct. R. Brown (1965, p. 647) says that the self is the person of particular interest to us, and describes it as “a cognitive construction of the human organism.” Studies take two main approaches, with some seeking to determine the influence of the self-beliefs and others trying to determine how humans acquire an aspect of self. Self-confidence or confidence is the aspect of self that is of particular interest to this study.

This section reviews literature related to the subject of confidence in one’s ability to do and learn mathematics. When studying self-beliefs one encounters the problem of definitions, and the interchangeable use made of different aspects of self. The section therefore begins by examining some of the definitions in use (Section 2.3.1) and continues by comparing confidence with other self-beliefs such as self-concept and self-efficacy. It further describes some of the instruments used in measuring different self-beliefs, with a particular focus on the ATMI. Section 2.3.2 reviews literature relating to relationships and differences between confidence and the variables which form part of this study, namely achievement, year level, ability grouping and strand of mathematics studied.

Over time, confidence has been identified as an attitude (Fennema & Sherman, 1976, 1977; Reyes, 1984; Reyes & Stanic, 1988; Tapia & Marsh, 2004), an affective variable (Cretchley, 2008), a belief (Benabou & Tirole, 2002; Druckman & Bjork, 1994) and a personality trait (Kleitman & Stankov, 2007; Stankov, 1999; Stankov et al., 2015). This section reviews some of the commonly used definitions of the term, with particular reference to confidence in mathematics.

The early research that showed an interest in confidence in relation to mathematics was studying attitudes towards mathematics. Much of this work was interested in determining whether attitudes such as confidence played a role in the different achievement and participation rates of males and females in mathematics. The focus was on measuring levels of confidence and their link to these factors, not on determining how confidence was acquired. The focus of this study is therefore on measuring students’ confidence in their ability to do and learn mathematics and determining whether it is related to their perceptions of the classroom learning environment.

Amongst the first to include a confidence scale in their instruments to measure attitudes were Fennema and Sherman (1976). They developed one of the early multi-faceted instruments used to measure attitudes, the Fennema-Sherman Mathematics Attitude Scale (FSMAS), for use in their groundbreaking research on gender differences in mathematics. It included a Confidence In Learning Mathematics scale which was

> intended to measure confidence in one's ability to learn and to perform well on mathematical tasks. The scale is not intended to measure anxiety or mental confusion, interest, enjoyment, or zest in problem solving. (Fennema & Sherman, 1976, p. 326)

Reyes drew on the work of Shavelson, Hubner, and Stanton (1976) when she defined confidence in learning mathematics as “a particular component of self-concept that is specific to mathematics” (Reyes, 1984, p. 559). For her, confidence “has to do with how sure a person is of being able to learn new topics in mathematics, perform well in mathematics class, and do well on mathematics tests” (Reyes, 1984, p. 560).

Druckman & Bjork (1994) make a distinction between specific and global confidence, saying that confidence can be viewed as a belief in one’s ability to perform specific tasks or activities, or as a global trait that accounts for beliefs in overall performance and optimism. Confidence as used in this study refers to self-beliefs about one’s ability to do and learn mathematics; it is a global concept in that it accounts for students’ beliefs in their overall performance in mathematics.

2.3.1 Confidence and Other Self-Perception Constructs

The lack of consistency in the usage of terms for self-perception is problematic when trying to compare the results of different studies (Burton, 2004; B. M. Byrne, 1996; Cramer, Neal, & Brodsky, 2009; Cretchley, 2008; Di Martino & Zan, 2010, 2011, 2015; Hattie, 1992; Leder & Forgasz, 2006; Liljedahl & Hannula, 2016; Marsh, 1990a; Marsh et al., 2019; Pajares, 1996). The terms most often used interchangeably with confidence are self-concept and self-efficacy. Section 2.3.1.1 examines self-concept and Section 2.3.1.2 self-efficacy. Section 2.3.1.3 outlines the differences and similarities between these self-beliefs by describing the different ways in which they are measured, and concludes by considering the division of self-perceptions into two categories: self-concept-like and self-efficacy-like. Finally, Section 2.3.1.4 describes the Attitudes Towards Mathematics Inventory (ATMI), the instrument used to measure confidence in this study.

2.3.1.1 Self-Concept

In order to provide a universal definition of the term self-concept, Shavelson et al. (1976) proposed a hierarchical model, with general self-concept at the top and academic, social, emotional and physical self-concepts at the next level down. They described self-concept as a hypothetical construct that could be useful in predicting how a person would behave, but was not an entity within the person. Academic self-concept was further divided into specific subjects: mathematics self-concept, science self-concept, etc. Thus, self-concept was defined as an individual's perception of themselves formed through their experience and interpretation of their environment and influenced by the evaluations of significant others.

Seven features were identified as being critical to the construct of self-concept: "organized, multifaceted, hierarchical, stable, developmental, evaluative, differentiable" (Shavelson et al., 1976, p. 411). Song and Hattie (1984) proposed a modification to the Shavelson model by dividing academic self-concept into three factors: achievement self-concept, ability self-concept and classroom self-concept. Later, the Marsh/Shavelson model proposed two higher-order academic factors: math/academic and verbal/academic (Marsh, 1990b). This hierarchical model still informs much recent research into academic self-concept (Brunner et al., 2010; Jansen, Schroeders, & Lüdtke, 2013; Jansen, Schroeders, Lüdtke, & Marsh, 2019; Waddington, 2019).

2.3.1.2 Self-Efficacy

The term self-efficacy has been in use since the 1970s, when Bandura (1977) described its role in behavioural change. For Bandura (1997, p. 382), "Perceived self-efficacy refers to belief in one's agentive capabilities, that one can produce given levels of attainment". He proposed four sources of information from which self-efficacy beliefs are constructed. These are “enactive mastery experiences”, “vicarious experiences”, “verbal persuasion” and “physiological and affective states” (Bandura, 1977, 1997), and deemed enactive mastery to be the most influential of these, as it is “the most authentic evidence of whether one can muster whatever it takes to succeed” (Bandura, 1997, p. 80). He also identified four processes through which efficacy beliefs produce their effect: cognitive, motivational, affective and selective processes. He differentiated between outcome expectancy and efficacy expectancy, with outcome expectancy being “a person’s estimate that a given behaviour will lead to a certain outcome” and efficacy expectancy being “the conviction that one can successfully execute the behaviour required to produce the outcomes” (Bandura, 1977, p. 193). For Bandura (1986), self-efficacy is not concerned with an individual’s skills but with what they can achieve with that set of skills.

2.3.1.3 Differences in Measuring Self-Beliefs

One of the obvious differences between self-beliefs relates to their levels of generality and how they are measured. Self-concept provides a broader range of information about self-perceptions of one’s abilities to do and learn mathematics than self-efficacy, which refers to self-beliefs about one’s ability to perform specific mathematical tasks (Bong & Skaalvik, 2003; Cretchley, 2008; Pajares, 1996; Stankov et al., 2015). The differences in specificity become clearer when one examines some of the instruments used to measure these constructs. This section therefore looks at instruments which have been developed to measure self-perceptions such as confidence, self-concept, self-efficacy and academic efficacy.

A number of instruments have been developed to measure confidence. The first and most significant of these was the FSMAS, which was developed by Fennema and Sherman (1976) to assess attitudes towards mathematics as part of their research in order “to gain more information concerning females’ learning of mathematics as well as information concerning variables related to the selection of mathematics courses” (Fennema & Sherman, 1976, p. 325). It became the most widely used attitude scale for the next thirty years (Tapia & Marsh, 2004). This instrument was not used in this study because it is 45 years old, making its language dated, and because it is very long, taking about 45 minutes to complete (Mulhern & Rae, 1998).

Other, similar instruments have been developed for use in research, in particular exploring the impact of the use of technology on mathematics. Galbraith and Haines (1998) developed an instrument to measure students' attitudes towards mathematics and technology in a computer learning environment. It included a scale called mathematics confidence. This included items such as "mathematics is a subject in which I get value for effort", "I can get good results in mathematics" and "I have a lot of confidence when it comes to mathematics" (Galbraith & Haines, 1998, p. 281). Pierce, Stacey, and Barkatsas (2007) developed the Mathematics and Technology Scale (MATS) for use in their research into students' attitudes to learning mathematics with technology. It only took about 10 minutes to complete and included items similar to those in the other instruments mentioned above, e.g., "I have a good mathematical mind" and " I am good at mathematics" (Pierce et al., 2007, p. 299). These instruments were not used in this study because they were designed for use in computer learning environments.

Kleitman and Stankov (2007) and Stankov et al. (2015) proposed two ways of measuring confidence. The first was using self-report questionnaires similar to those described above to assess a person's beliefs in their ability to accomplish different tasks. This is a general measure which reflects self-belief about one's abilities to do something in a particular field (e.g., mathematics) without the need for verification. The second method involved making a judgement about one's likelihood of answering a question correctly immediately after having answered it (i.e., after a cognitive act). While both types of measures capture personality traits and cognitive aspects of confidence, the difference between them is in their emphasis, with the second method closely resembling that recommended by Bandura (1977) for measuring self-efficacy.

Hattie (1992) proposed that the best of the large number of measures of self-concept available was the Marsh Self-Description Questionnaire (SDQ) (Marsh, Relich, & Smith, 1983), which is based on the Marsh/Shavelson model of self-concept. The mathematics self-concept scale includes item such as "I am interested in mathematics", "I am good at mathematics" and "I learn things quickly in mathematics" (Marsh et al., 1983, p. 179).

The Programme for International Student Assessment (PISA 2003) included five self-concept scales (J. Lee, 2009; Morony, Kleitman, Lee, & Stankov, 2013) which included items such as “I am confident when it comes to mathematics”, “I learn mathematics easily” and “I am good at solving mathematics problems”. These are very similar to what are called confidence items in the Trends in International Mathematics and Science Study (TIMSS 2007) (Thomson, Wernett, Underwood, & Nicholas, 2008), which include “I usually do well in mathematics”, “Mathematics is harder for me than for many of my classmates”, “I am just not good at mathematics” and “I learn things quickly in mathematics”. The similarities in the items used would suggest that they are in fact measuring the same construct, despite using different names for it.

Self-efficacy is measured using task-specific questionnaires. Nielsen and Moore (2003) developed the Mathematics Self-efficacy Scale, which asked students to estimate their own mathematical ability by stating how confident they were that they could complete particular mathematics tasks such as solving a simultaneous equation, calculating the surface area and volume of a solid or working with fractions and decimals. The focus was on performance capabilities rather than psychological characteristics. “A self-efficacy assessment, therefore, includes both an affirmation of a capability level and the strength of that belief” (Bandura, 1997, p. 382), and is highly content and context specific. The question is usually answered before attempting to solve the problem.

Another self-perception prevalent in the literature pertaining to classroom learning environments is academic efficacy. This is usually measured using the Academic Efficacy scale from the Patterns of Adaptive Learning Scales (PALS) developed by Midgley et al. (2000). The Academic Efficacy scale “refers to students’ perceptions of their competence to do their class work” (Midgley et al., 2000, p. 19). It includes items such as “I'm certain I can master the skills taught in class this year” and “Even if the work is hard, I can learn it”.

Marsh et al. (2019) categorised self-belief constructs as being either self-concept-like or self-efficacy-like. Self-concept-like instruments ask general questions about students’ perceptions of their mathematical ability and performance, while self-efficacy-like instruments ask about their confidence in their ability to correctly answer a specific question. Marsh et al. (2019) conducted a secondary data analysis of a large-scale longitudinal study in Bavaria and found a strong correlation (.9) between three self-concept-like constructs: self-concept, outcome expectancy and academic self-efficacy. They also found two self-efficacy measures (test self-efficacy and functional self-efficacy) to be distinct from the self-concept-like constructs but not from each other. Based on their classifications, the types of items used in instruments to measure confidence, self-concept and academic efficacy as described above would fit into the self-concept-like category. Thus studies using these constructs could reasonably be compared.

This section has described some of the differences between self-belief constructs in common use today. It has identified self-beliefs as either self-concept-like or self-efficacy-like. Confidence, as used in the study reported in this thesis, is a self-concept-like construct which refers to self-beliefs about one's ability to do and learn mathematics; it is global in that it accounts for students' beliefs in their overall performance in a particular strand of mathematics rather than their ability to answer specific mathematics questions.

2.3.1.4 Using the Attitudes Towards Mathematics Inventory (ATMI) to Measure Confidence

This section will describe the Attitudes Towards Mathematics Inventory (ATMI), which was used as the basis for the questionnaire in the study reported in this thesis. It reviews research which has made use of this instrument.

Late last century, Tapia (1996) began the development of an instrument to measure attitudes that also included a confidence scale. It was the Attitudes Towards Mathematics Inventory (ATMI), a modified version of which was used in this study. The ATMI was developed because of the need for a shorter scale than the FSMAS which would withstand factor analysis and "tap important dimensions of attitude" (Tapia & Marsh, 2004, p. 17).

The final version of the ATMI has four scales (Self-Confidence, Value, Enjoyment and Motivation) and a total of 40 items. The Self-Confidence scale, comprising 15 items, was designed to assess "the confidence and self-concept of students with respect to their performance in mathematics" (Tapia & Marsh, 2004, p. 17). The Value scale has 10 items and measures "students' beliefs on the usefulness, relevance and worth of mathematics to their lives" (Tapia & Marsh, 2004, p. 17). The Enjoyment scale, with 10 items, measures "the degree to which students enjoy working on mathematics" (Tapia & Marsh, 2004, p. 17). The Motivation scale, with five items, measures students' "interest in mathematics and their desire to pursue further studies in mathematics" (Tapia & Marsh, 2004, p. 17).

In order to determine its suitability for use with other groups of students, Tapia and Marsh (2005) administered the ATMI to 134 college-aged American students. The purpose of this study was to determine whether the ATMI, which had previously been found reliable and valid for Hispanic students (Tapia, 1996; Tapia & Marsh, 2000, 2002, 2004), would produce similar statistical properties when used with college-aged American students. The sample made up of 80% Caucasian and 20% African-American students, produced similar results to the previous study with Hispanic students. The Cronbach coefficients for each of the factors ranged from .96 for Self-Confidence to .87 for Motivation. The results of the confirmatory factor analysis confirmed "the four-factor structure as the best simple fit for the items, which were found in the original study with secondary students" (Tapia & Marsh, 2005, p. 275).

The ATMI has been used to explore the relationships between attitudes and mathematics achievement (Davadas & Lay, 2020; Ke, 2008; Yee, 2010), attitudes and gender differences (Asante, 2010, 2012; Davadas & Lay, 2020; Lin & Huang, 2014) and attitudes and school types (Davadas & Lay, 2020; Khan & Rodrigues, 2012). Other researchers (Afari, 2013; Khine & Afari, 2014; Khun-inkeeree & Omar-fauzee, 2020; Lim & Chapman, 2013; Lin & Huang, 2014; Majeed, Darmawan, & Lynch, 2013; Primi, Bacherini, Beccari, & Donati, 2020) have created shorter versions and conducted factor analysis to confirm the validity and reliability of the ATMI. More information about the ATMI as used in this study is given in Chapter 3.

The next section reviews literature relating to the association between confidence and achievement, and differences in confidence due to year level, ability grouping and strand of mathematics studied.

2.3.2 The Relationship Between Confidence and Achievement, Year Level, Ability Grouping and Strand of Mathematics Studied

The study reported in this thesis explores whether relationships exist between confidence and achievement and between confidence and classroom learning environments. It also examines differences in confidence based on three variables: year level, ability grouping and strand of mathematics being studied. Section 2.3.2.1 reports on studies which have examined the relationship between confidence and achievement. The differences in confidence levels due to year level, ability grouping and strand of mathematics studied are reported in Sections 2.3.2.2, 2.3.2.3 and 2.3.2.4 respectively. Research relating to classroom learning environments and confidence is discussed in Section 2.4.

2.3.2.1 Confidence and Achievement

This section provides information about some of the problems encountered in research relating to the relationship between confidence and achievement. These difficulties primarily relate to issues with defining both terms. The section also examines research done in this area, with particular reference to studies concerning confidence and achievement.

While common sense would suggest that there must be a relationship between achievement and self-beliefs such as confidence, past research has been inconclusive (Hattie, 1992; Huang, 2011; Ma & Kishor, 1997a). According to Hattie (1992, p. 200) the general opinion "is that a moderate and positive association exists between self-concept and measures of performance and achievement". However, when looking at the research literature on the topic, he concluded that "the relationship is neither precise nor clear" (Hattie, 1992, p. 200). This lack of clarity may be due to variations in how self-concept and other self-beliefs such as self-concept, self-efficacy and confidence are defined and measured. Huang (2011, p. 505) reported that the "strength, significance and generality" of the relationship between self-concept and academic achievement has been inconclusive. He conducted a mega-analysis of research examining the relationship between self-concept and academic achievement using a longitudinal sample, and concluded that there is a stronger relationship between the two when self-concept is measured at a subject-specific level rather than at a global level. This result is supported by other research (Marsh, 1990a; Marsh, Byrne, & Shavelson, 1988; Marsh, Byrne, & Yeung, 1999; Pajares & Schunk, 2001). Huang (2011, p. 526) also concluded that the "findings indicate high self-concept leads to high academic achievement and vice versa".

In addition to problems with defining and using the self-belief terminology, there have also been problems with how achievement is measured. B. M. Byrne (1996) reported that prior to 1980, as well as inconsistencies in measuring self-concept, there were problems with what was meant by achievement, with measures ranging from subject grades to standardised test scores, self-reported grades, teacher ratings and school board examinations.

Research using the confidence scale of the FSMAS, a self-concept-like scale similar to the one used in this study, has consistently reported positive correlations between confidence and achievement. This included cases where achievement was measured using standardised tests (Fennema & Sherman, 1977, 1978; Tartre & Fennema, 1995) and teacher-set tests (Nurmi et al., 2003). Yee (2010) found positive correlations between confidence and achievement when confidence was measured using the ATMI and a school board examination was used to measure achievement. International studies such as the TIMSS (Thomson et al., 2017) and the PISA (J. Lee, 2009) have also found positive correlations between achievement, measured using standardised tests, and confidence, measured on self-concept-like scales.

It is therefore reasonable to conclude that there is a positive relationship between confidence and achievement, but that its exact nature is unclear. There are three possibilities: that an increase in confidence leads to an increase in the level of achievement; that an increase in achievement level results in an increase in confidence, or that the relationship is reciprocal. A growing number of researchers (Arens et al., 2017; Ganley & Lubienski, 2016; Huang, 2011; Marsh & Martin, 2011; Niepel, Brunner, & Preckel, 2014; Schöber, Schütte, Köller, McElvany, & Gebauer, 2018; Valentine, DuBois, & Cooper, 2004) would argue that there is a reciprocal relationship between confidence and achievement and that it is important for teachers to work on enhancing both at the same time. This is not an easy thing to do because, as pointed out by Marsh et al. (2017, p. 277), teachers “need to teach academic skills, but also to reinforce positive self-beliefs and link the two in a way that is consistent with their reciprocal relations”.

2.3.2.2 Confidence and Year Level

As well as examining the relationship between confidence and achievement, this study also investigates whether students exhibit different levels of confidence at different stages of their schooling. This section therefore reviews past studies that have examined changes in students’ self-perception as they progress through school, in particular studies which have examined changes at the critical time of transitioning to high school.

Numerous studies have reported a change in students’ perceptions of their abilities as they transition to high school (Midgley, Feldlaufer, & Eccles, 1989; Seidman, Allen, Aber, Mitchell, & Feinman, 1994; Speering & Rennie, 1996; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991; Wouters, De Fraine, Colpin, Van Damme, & Verschueren, 2012; Yeung, 2011). However, in a review of literature pertaining to students’ self-beliefs at the transition to high school, Eccles and Midgley (1989) found a lack of consistency in the results, with some studies showing no difference in students’ confidence in their abilities at transition.

An analysis of the TIMSS 2015 data (Thomson et al., 2017) showed that Year 8 students in Australia were less confident in their mathematical ability than their Year 4 counterparts. This confirmed similar findings using the 2007 TIMSS data (Thomson et al., 2008). A large Finnish study (Nurmi et al., 2003), “Understanding and Self-Confidence in Mathematics”, found that students in higher year levels were less confident than those in lower year levels. They surveyed 3,057 students in 150 schools and measured confidence using items from the confidence scale of the FSMAS. Another, longitudinal study (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002), involving 761 students across Years 1 to 12 found that the students’ self-perceptions of competence declined as they grew older. This decline occurred in all three domains included in their study: mathematics, language arts and sport.

Studies which report a change in students' confidence at transition report that it declines. The research reported in this study also finds that students become less confident in their ability to do and learn mathematics as they progress through school, although the differences are not statistically significant.

2.3.2.3 Confidence and Ability Grouping

The third variable investigated in the study reported in this thesis is ability grouping: i.e., whether students in different ability groups are more or less confident in their ability to do and learn mathematics. There has been much written about the benefits of dividing students into groups on the basis of their ability, a procedure sometimes called tracking, streaming or setting (Belfi, Goos, De Fraine, & Van Damme, 2011; Hattie, 2002; Herrmann, Schmidt, Kessels, & Preckel, 2016; Houtte, Demanet, & Stevens, 2012; Ireson & Hallam, 2009). Generally, it has been found that there is a positive correlation between self-concept-like perceptions and ability grouping, but as Francis et al. (2017, p. 102) point out,

> we see a clear relationship between set level and self confidence in both the setted subjects concerned and in learning more broadly, compounding the implication in our findings that the labelling associated with allocation to sets triggers a self-fulfilling prophecy in relation to pupils' perceptions of the subject concerned, and their relationship to it.

In other words, past research has found that students in higher-ability mathematics classes may be more confident in their ability to do and learn mathematics by virtue of having been placed in that particular group.

2.3.2.4 Confidence and Strand of Mathematics Being Studied

The fourth variable investigated in this study is whether student confidence differs depending on the strand (or area) of mathematics studied. Anecdotally, students have told me that they feel more confident in their ability to do and learn one strand than another: e.g., Statistics vs Algebra. In one study, involving 1,314 third grade and 1,412 fifth grade students, Vanayan, White, Yuen, and Teper (1997) found that students were able to identify clear differences in beliefs and attitudes towards different strands of mathematics. In developing her hierarchy of definitions for the three self-beliefs – self-concept, self-confidence and self-efficacy – Cretchley (2008, p. 152) said that

> Self-confidence (usually termed just confidence) refers to self-beliefs about abilities to do and learn mathematics in some context, not necessarily generally. Hence a learner may be confident within one area of mathematics, but perhaps not another.

However, little research has examined whether this is in fact the case. What in fact students may be referring to is their preference for one strand over another based on their enjoyment, the perceived difficulty or the perceived usefulness of that strand.

This section has pointed out some the differences and similarities between the numerous self-beliefs used in research. It has concluded that there are two main types of self-beliefs, self-concept-like and self-efficacy-like constructs, with confidence, the focus of this study, fitting in the first category. Information about instruments used to measure self-beliefs has been included, with a focus on the ATMI, which is used in this study. The literature reviewed in this section generally supports the view that there is a positive relationship between achievement and confidence, although there are inconsistencies in research results. There is also evidence to support the view that there are differences in confidence levels based on membership of a particular year level or ability group, with those in higher year levels being less confident and those in higher ability groupings being more confident. There is not enough evidence to determine whether confidence differs due to the strand of mathematics being studied. The next section will examine classroom learning environment research in detail.

2.4 Classroom Learning Environment Research

The main focus of this study is on exploring the relationships between female students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. This section provides a detailed review of literature related to the field of learning environment research.

Fraser (1998a, p. 3) defines the classroom learning environment as “the social, psychological, and pedagogical contexts in which learning occurs and which affect achievement and attitudes”. It is a human environment which includes teachers and students. It has been shown to be one of the leading factors in academic wellbeing (Walberg, 1984). Fraser (2001) believes that students are well placed to assess the classroom learning environment because they have up to 20,000 hours of experience of a variety of classrooms.

Numerous studies have provided evidence to suggest that the quality of the classroom learning environment is a significant determinant of student learning outcomes (Chionh & Fraser, 2009; Fraser, 1998b; Fraser & Fisher, 1982; Ogbuehi & Fraser, 2007). Some of these studies “have shown that student perceptions account for appreciable amounts of variance in learning outcomes, often beyond that attributable to background student characteristics” (Fraser, 2012b, p. 1281). Recently, Fraser (2019, p. 2) stated that

> What the past 30 years of learning environment research has convincingly established is that the subtle construct of learning environment can be reliably and validly measured (typically through the eyes of the participants themselves) and that attention to creating positive classroom environments is likely to pay off in terms of improving student outcomes.

This section reviews literature related to the almost 100 years of interest in this topic. Firstly, it reviews the history of classroom learning research from its inception in the 1930s with the work of Lewin and Murray (Section 2.4.1). Secondly, it reviews some of the many classroom learning environment instruments that have been developed for use in such research (Section 2.4.2). Finally, it reviews some of the vast quantity of research that has been conducted using classroom learning environment instruments (Section 2.4.3).

2.4.1 Historical Background of Classroom Learning Environment Research

The study of the field of classroom learning environments had its beginnings in the ideas of social psychologists Kurt Lewin and Henry Murray, both of whom postulated that an individual’s behaviour depends on both the state of the environment and the personality of the individual. Lewin (1936) used the formula $B = f(P, E)$, where B is behaviour, P is personality and E the environment, to describe what he called the whole psychological situation. Lewin stressed the need to recognise this relationship when attempting to describe any behaviour, whether physical or mental. Meanwhile, Murray (1938) built on Lewin’s work to develop a needs-press model, in which he described personal needs as personality traits which influence motivation and environmental press as the external factors that work for or against the personal needs of the individual. These breakthrough ideas were reflected in the work that followed.

In the 1960s Moos and Houts (1968), positing that little attention had been paid to the environmental dimensions of personality, suggested the need to develop ways of assessing environments. Initially instruments were developed for use in psychiatric hospitals, but their use was soon expanded to numerous other social settings such as prisons and schools. As a result of his work on personality, Moos (1973) developed a scheme of classifying human environments. He identified nine environments (including schools, hospitals and prisons) and developed three dimensions to enable the classification and sorting of various components of these human environments. The Relationship Dimensions, which are similar across all nine environments, assess the extent to which individuals help and support each other and are involved in the environment. In a classroom they include factors such as cohesiveness, involvement and teacher support. The Personal Development Dimensions assess the way in which personal development and self-enhancement occur in the particular environment. The nature of these dimensions varies depending on the environment’s basic purpose and goals; in educational environments they assess factors such as task orientation, difficulty and competitiveness. The System Maintenance and System Change Dimensions are similar across all nine environments identified by Moos. The basic dimensions are diversity, formality, organisation, rule clarity and equity. A dimension of innovation is also identified in educational and small group environments.

In the 1960s and 1970s, the work of Walberg and Anderson changed the direction of classroom learning environment research. While previous studies had treated the class as an entity, their work focussed on the individual, and they sought to establish whether there was a relationship between the student's individual satisfaction with the classroom learning environment and their learning (Anderson, 1970; Walberg, 1969; Walberg & Anderson, 1968a, 1968b). Their work set the stage for the classroom learning environment research that was to come. During this time, the first learning environment questionnaires – the Learning Environment Inventory (LEI) (Walberg & Anderson, 1968b) and the Classroom Environment Scale (CES) (Trickett & Moos, 1973) – were developed. These established the pattern for future questionnaires by using Moos social climate scales.

Trickett and Moos (1973) asserted that the instrument they developed, the Classroom Environment Scale (CES), would enable them and others to study classroom learning environments in more depth and to determine the role of all participants – students and teachers – in the development of the classroom learning environment. They also stated that it could be used to help teachers reflect on their behaviour and make changes to improve their students' outcomes.

During the 1980s, the importance of the school environment as a factor in the effectiveness of schools was recognised with the development of instruments such as the School-Level Environment Questionnaire (SLEQ) (Rentoul & Fraser, 1983). This period also saw the development and use of actual and preferred forms of the learning environment instruments. One of the first studies in this area was conducted by Fisher and Fraser (1983). They used scales from the ICEQ and the CES to compare perceptions on three forms of those surveys: student actual, student preferred and teacher actual. They found that students prefer a more positive environment than that which they perceive and that teachers perceive the classroom learning environment more favourably than the students do. In another study, D. B. Byrne, Hattie, and Fraser (1986) used short forms of the MCI, CES and ICEQ to obtain data about students' preferred classroom learning environments and the Quality of School Life questionnaire to ascertain their perceptions of their actual environments. They reported that students preferred a more positive environment on a number of dimensions than was actually present, and that there were changes in preference based on age. This led them to conclude that "The differences in preferred classroom environments over ages helps to explain why some researchers have noted a 'decline' in classroom environment as students progress through school" (D. B. Byrne et al., 1986, p. 17).

By the early 1990s a new wave of research into the classroom learning environment began with two programmes, one in the Netherlands and the other in Australia. In the Netherlands, Wubbels (1993) developed the Questionnaire on Teacher Interaction (QTI), which focussed on the interactions between students and teachers. His research showed that "interpersonal teacher behaviour is an important aspect of the learning environment. It is strongly related to student outcomes" (Wubbels, 1993, p. 6).

In Australia, Fraser's research focused on student centred classrooms through the use of the ICEQ (Burden & Fraser, 1993) and subsequently the development and validation of a number of other classroom learning environment instruments, including the Science Laboratory Environment Inventory (SLEI) (Fraser, Giddings, & McRobbie, 1992, 1995) the Constructivist Learning Environment Survey (CLES) (Nix, Fraser, & Ledbetter, 2005) and the WIHIC (Aldridge et al., 1999; Fraser, et al., 1996). This research established classroom learning environment instruments as reliable and valid ways of measuring students' perceptions of the classroom learning environment. It provided evidence "that student perceptions account for appreciable amounts of variance in learning outcomes, often beyond that attributable to background student characteristics" (Fraser, 2012b, p. 1218).

The increase in the number of questionnaires available for collecting data has led to a wide variety of types of classroom learning environment research. Fraser (2012b) has identified many of these, the main ones being: associations between student outcomes and environment (Afari, Aldridge, Fraser, & Khine, 2013; Aldridge, Afari, et al., 2012; Dorman, 2001; Dorman & Adams, 2004; Dorman, Adams, & Ferguson, 2003; Dorman et al., 2006; J. M. Ferguson & Dorman, 2001; Helding & Fraser, 2013; Martin-Dunlop & Fraser, 2007; Ogbuehi & Fraser, 2007; Wolf & Fraser, 2008); the evaluation of educational innovations (Aldridge & Fraser, 2011; Martin-Dunlop & Fraser, 2007; Nix et al., 2005); and teachers' attempts to improve their classroom and school environments (Bell & Aldridge, 2014; Yarrow, Millwater, & Fraser, 1997). Other types of research using learning environment instruments have included: explorations of the differences between perceptions of actual and preferred environments (Aldridge, Fraser, Bell, & Dorman, 2012); research combining quantitative and qualitative methods (Tobin & Fraser, 1998); cross-national studies (Aldridge et al., 1999; Aldridge, Fraser, & Ntuli, 2009; MacLeod & Fraser, 2010); examinations of the transition from primary to high school (J. M. Ferguson & Dorman, 2001); and studies of teacher education (Aldridge et al., 2009; Martin-Dunlop & Fraser, 2007).

This section has provided a brief overview of the history and development of research in the field of classroom learning environments. The next section reviews some important instruments that have been developed over the years to aid in this research.

2.4.2 Classroom Learning Environment Instruments

There are many questionnaires available for researchers to use in the field of classroom learning environment research. They provide a convenient way of collecting data and have been developed to address particular features of the classroom learning environment, such as relationships between the teacher and the students, student participation in the class, the nature of teacher support in the class, and student cooperation. Many of the questionnaires have been translated into several languages, including Chinese, IsiZulu and Arabic (Aldridge, Fraser & Huang, 1999; Aldridge, Fraser & Ntuli, 2009, MacLeod & Fraser, 2010). This section describes eight historically important classroom learning environment instruments that were developed based on Moos' three dimensions of human environments. It focusses on the ones most relevant to the issues covered in this study: the Learning Environment Inventory (LEI) (Section 2.4.2.1); the Classroom Environment Survey (CES) (Section 2.4.2.2); the Individualised Classroom Environment Questionnaire (ICEQ) (Section 2.4.2.3); the My Class Inventory (MCI) (Section 2.4.2.4); the College and University Classroom Environment Inventory (CUCEI) (Section 2.4.2.5); the Questionnaire on Teacher Interaction (QTI) (Section 2.4.2.6); the CLES (Section 2.4.2.7); and the WIHIC (Section 2.4.2.8).

2.4.2.1 The Learning Environment Inventory (LEI)

The LEI was one of the first questionnaires developed in the field of classroom environment research. It was developed by Walberg and Anderson as part of the Harvard Project Physics program in the 1960s (Walberg & Anderson, 1968a) for use with high school science classes. The final version contains 105 statements divided into 15 scales, with 7 items per scale and 4 possible responses ranging from Strongly

Disagree to Strongly Agree. The scales are Cohesiveness, Friction, Favouritism, Cliqueness, Satisfaction, Apathy, Speed, Difficulty, Competitiveness, Diversity, Formality, Material Environment, Goal Direction, Disorganisation and Democracy. The LEI has two distinct uses: “to assess the perceptions of an individual student, or to gauge the learning environment of the class as a group” (Fraser, Anderson, & Walberg, 1982, p. 4). While this instrument was designed for the classrooms of the 60s and 70s, it is clear from the use of scales such as Cohesiveness and Democracy that social relationships between students and fairness of the classroom environment were recognised as important aspects of learning environments from the beginning. Fraser et al. (1982) assessed the instrument as valid and reliable based on internal consistency reliability and discriminant validity methods. Although its factor structure has not been established, it has been widely used to research students’ perceptions of their classroom learning environments and a range of outcomes (Haertel, Walberg, & Haertel, 1981; Walberg, 1976; Walberg & Anderson, 1968b).

The LEI was not chosen for use in this study because the language was considered to be dated. With 105 items, it was also too long, especially as students were also required to complete a questionnaire relating to their perceptions of their confidence in their ability to do and learn mathematics.

2.4.2.2 The Classroom Environment Scale (CES)

The CES, one of the earliest questionnaires, was developed by Trickett and Moos (1973) as part of a programme of research involving a variety of human environments including hospitals, prisons and universities. It asks teachers and students to comment on various aspects of a class. The original version consisted of 242 items divided between 13 psychosocial dimensions of the classroom. It was trialled in 26 high schools. The final version has 9 scales: Involvement, Affiliation, Teacher Support, Task Orientation, Competition, Order and Organisation, Rule Clarity, Teacher Control and Innovation, with 10 items per scale. Responses are either True or False. Trickett and Moos (1973) purported that this instrument would allow more in-depth study of classroom learning environments by collecting data from both students and teachers. The validity and reliability of the scale has been reported in a number of studies (Fisher & Fraser, 1983; Trickett & Moos, 1973).

The CES was not chosen for use in this study because like the LEI, its language is dated and it was designed for older-style teacher-centred classrooms. Another problem is that it uses a true-false response scale, which makes it difficult to gain an accurate assessment of students' perceptions of the classroom learning environment.

2.4.2.3 The Individualized Classroom Environment Questionnaire (ICEQ)

Rentoul and Fraser (1979) developed the ICEQ after extensive interviews of teachers and secondary school students in Australia. It measures classroom learning environment perceptions along dimensions that differentiate individualised classes involving enquiry-based or open approaches from more conventional ones. It has different forms to measure students' or teachers' perceptions of actual or preferred classrooms. The final version has 5 scales with 10 items per scale. The scales are Personalisation, Participation, Independence, Investigation and Differentiation. Each item is answered with one of five options, ranging from Almost Never to Very Often. This was one of the first instruments to accommodate the changes to classroom methodologies that were beginning to take place in the late 1970s. Its validity and reliability have been reported by a number of studies (Asghard, 1994; Fraser, 1980; Fraser & Butts, 1982; Rentoul & Fraser, 1979).

As the ICEQ was developed for use in inquiry-based science classes and the mathematics classes involved in this study did not use inquiry-based learning, it was decided that it was not an appropriate choice for use in this study.

2.4.2.4 The My Class Inventory (MCI)

The MCI began as a simplified version of the LEI and was originally developed for use with primary school students (Fraser et al., 1982). It was further simplified by Fisher and Fraser (1981) into a 38-item instrument. It is different from the LEI in four ways. There are fewer scales and items – 5 rather than 15 scales and 38 rather than 105 items – and the responses are on a 2-point Yes/No scale rather than the 4-point scale of the LEI, although this was increased to 3 (Seldom, Sometimes, Most of the Time) by Goh and Fraser (1998). The language of each item has also been simplified to be more appropriate for children aged 8–12 years. Finally, the answers are placed on the question sheet rather than on a separate answer sheet. The five scales used are Cohesiveness, Friction, Satisfaction, Difficulty and Competitiveness, with six to nine items per scale. Goh and Fraser (1998) also added a Task Orientation scale. A 25-item short form was developed by Fraser and O'Brien (1985), which has 5 scales with fewer items per scale. The MCI has been used and validated in a number of studies (Houston, Fraser, & Ledbetter, 2008; Majeed, Fraser, & Aldridge, 2002).

The MCI was developed for use with primary-school students; consequently, as this study was focussed on middle-school students, it was not chosen for use. Its use of a two-point Yes/No response scale also rendered it unsuitable.

2.4.2.5 The College and University Classroom Environment Inventory (CUCEI)

To redress the lack of research on classroom learning environments being done in higher education classrooms, Fraser and Treagust (1986) developed the CUCEI for use with small (<30) university classes. The final version has seven scales with seven items each. The scales are Personalisation, Involvement, Student Cohesiveness, Satisfaction, Task Orientation, Innovation and Individualisation. Three of these scales are in this study (Involvement, Student Cohesiveness and Task Orientation). Answers are selected from four choices ranging from Strongly Agree to Strongly Disagree. Although the CUCEI has been used validated in a number of studies (Fraser, 1991; Fraser & Treagust, 1986; Yarrow et al., 1997), its psychometric properties presented problems when used in a study involving computing classes in New Zealand (Logan, Crump, & Rennie, 2006).

The CUCEI was not chosen for use in this study because it was developed for use with students in higher education environments. It was also rejected because its validity and reliability have not been well established.

2.4.2.6 The Questionnaire on Teacher Interaction (QTI)

The QTI was developed by Wubbels (1993) to ascertain students’ perceptions of eight aspects of behaviour: Leadership, Helpful/Friendly, Understanding, Student responsibility/Freedom, Uncertain, Dissatisfied, Admonishing and Strict. The eight scales or aspects of behaviour are responded to using a five-point scale with responses ranging from Never to Always. In keeping with the classroom learning environment research that originated in the Netherlands, it focuses on the nature of the relationship between students and teachers. The QTI is a widely recognised learning environment survey that has been validated by numerous past researchers (Fisher, Henderson, & Fraser, 1995; Goh & Fraser, 1996; Wubbels & Levy, 1993).

Although student perceptions of this relationship were important to my study, they were not the only aspect of the classroom learning environment I wanted to explore. Despite its focus on the student-teacher relationship, the QTI was therefore not suitable for use in this study.

2.4.2.7 The Constructivist Learning Environment Survey (CLES)

The CLES was developed by Taylor, Fraser, and Fisher (1997) to assist teachers and researchers in determining to what extent constructivist practices are in use in a particular classroom. Constructivists believe that new knowledge is constructed from the learners' previous knowledge, beliefs and experiences of the world. This was the first instrument to group items belonging to the same scales together. The original version has 36 items divided between 5 scales (Personal Relevance, Uncertainty, Critical Voice, Shared Control and Student Negotiation), with 5 response alternatives ranging from Almost Never to Almost Always. A shorter form of the CLES was developed by Nix et al. (2005). It has 30 items and was developed to review the impact of an innovative teacher development programme. The researchers reported strong factorial validity and reliability. A 20-item version of the CLES was also developed for use in a study of beginning science teachers (Johnson & McClure, 2004). It retained the original five scales but reduced the number of items in each scale from six to four. Johnson and McClure (2004) reported that this version also exhibited strong factorial validity and reliability. The factorial validity and reliability of the CLES has been reported by numerous other studies (Kim, Fisher, & Fraser, 1999, 2000; Nix et al., 2005; Ogbuehi & Fraser, 2007; Sebela, Fraser, & Aldridge, 2003; Taylor et al., 1997).

Although constructivism underpinned the chosen framework used by one of the schools in the present study, the CLES was not selected as the learning environment instrument because this was not the main aspect of the classroom learning environment under consideration.

2.4.2.8 The What Is Happening in This Class? (WIHIC)

Dorman (2008) described the WIHIC as one of the foremost tools in classroom learning environment research. Like the CLES, it addresses the shift from teacher-centred to student-centred classrooms. A modified version was selected for use in this study; hence it will be described in more detail in this section. Firstly, I will cover its development; then I will provide details of some studies in which it has been used.

Development of the WIHIC

The WIHIC was developed because of the need for an instrument better suited to assessing differences in perceptions of the classroom learning environment held by different students or groups of students in the same class (Fraser et al., 1996). It includes scales which had previously been shown to be significant predictors of outcomes (Fraser, 1994) as well as new ones which addressed the then recent changes in classroom teaching. It measures high school students' perceptions of their classroom learning environment (Aldridge & Fraser, 1997).

Initially the WIHIC included nine scales: Student Cohesiveness, Teacher Support, Involvement, Autonomy/Independence, Investigation, Task Orientation, Cooperation, Equity and Understanding. There were 10 items for each scale, which were answered using a 5-point Likert-type scale from Almost Never to Almost Always. Personal and class forms were developed. The personal form elicits information about students' perceptions of their personal involvement in the class, while the class form measures their perception of the learning environment in the class. The WIHIC was originally administered in both class and personal forms, to 355 Australian Year 9 and 10 students in 17 classes. Following a series of item and factor analyses, the Autonomy/Independence and Understanding scales were removed and the WIHIC was refined to 54 items across 7 of the original scales. The results also showed statistically significant differences between the mean scores of the class and the personal forms of the questionnaires. Extensive interviews were conducted with 45 students to determine why they had given different answers on the class and personal forms. The data collected supported the view that the personal form assesses students' perceptions of their role in the class rather than asking them to provide perceptions of the class as a whole. It also supports the view that personal forms are required to assess "the different sub environments which might exist for different students or subgroups within the same class" (Fraser & Tobin, 1991, p. 291).

Further refinement of the WIHIC occurred during a cross-national study in 1999 that involved 1,081 Australian students and 1,879 Taiwanese students (using a translated version) in 50 junior high school science classes. (Aldridge et al., 1999). This study used a 70-item version of the questionnaire with 7 scales, each with 10 items. Fifty-six items survived the factor and item analysis, becoming the latest version of the WIHIC. Responses are given on a five-point Likert scale (Almost Never, Seldom, Sometimes, Often and Almost Always).

The seven scales of the WIHIC are: Student Cohesiveness, which refers to the extent to which students are supportive of each other; Teacher Support, which refers to the extent to which the teacher relates to the students; Involvement, which refers to the extent to which students participate in the class; Task Orientation, which refers to the extent to which students stay on task in lessons; Investigation, which refer to the extent to which the skills and processes of inquiry are used in the class; Cooperation, which refers to the extent to which students cooperate rather than compete; and Equity, which refers to the extent to which students are treated equally by the teacher. More detailed information about the WIHIC is included in Chapter 3.

Studies Reporting the Validity and Reliability of the WIHIC

The WIHIC has become one of the most used questionnaires developed (Fraser, 2012b), and its scales have also been used in other questionnaires such as the Constructivist-Oriented Learning Environment Survey (COLES) and the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI). This section describes some of the studies which have reported the validity and reliability of the WIHIC. These studies have been undertaken in a number of countries and languages and by different researchers.

Dorman (2003) undertook a comprehensive validation of the WIHIC in a sample including 3,980 high school students from Australia, Canada and the UK. Using confirmatory factor analysis (CFA), Dorman (2008) reported results which supported the seven-scale a priori structure. He later used a multitrait-multimethod approach with a confirmatory factor analysis framework to show the validity of the WIHIC construct. He administered the actual and preferred versions of the questionnaire to 978 Australian secondary school students and reported that the results of the analysis "demonstrate that the WIHIC has very sound structural characteristics", thus supporting "the view that the WIHIC is an important addition to the suite of high inference classroom environment instruments developed over the past 35 years" (Dorman, 2008, p. 190).

The WIHIC has also been used in a number of other cross-national studies in Taiwan and Australia and in Indonesia and Australia. As reported above, Aldridge et al. (1999) undertook a cross-national multiple research method study involving 1,081 Australian students and 1,879 Taiwanese students. The study made use of the 70-item WIHIC, which after principal components factor analysis resulted in the now more widely used 56-item version. The WIHIC was combined with "an eight-item scale to assess students' satisfaction in terms of enjoyment, interest, and how much they anticipated science classes according to a scale from the Test of Science Related Attitudes" (Aldridge et al., 1999, p. 50). The questionnaires were translated into Chinese for use in Taiwan. They were administered to 1,081 Year 8 and 9 general science students from 50 classes in 25 schools in Australia and 1,879 Year 7 to 9 students from 50 classes in 25 schools in Taiwan. The study showed the value of cross-national studies in helping to understand classroom learning environments in other countries as well as our own. It also supported the validity and reliability of all scales in both the English and Mandarin versions. Aldridge and Fraser (2000), working with 1,081 Australian students in 50 classes and 1,879 Taiwanese students in 50 classes, reported strong factorial validity and internal consistency reliability of the WIHIC. Fraser, Aldridge, and Adolphe (2010) used a modified version of the WIHIC in a cross-national study involving 567 Australian students in 18 classes and 594 Indonesian students in 78 secondary science classes. It was translated into Bahasa Indonesian and administered simultaneously in both countries. Principal components factor analysis with varimax rotation supported the validity of the modified structure of the WIHIC. While generally showing positive associations between classroom learning environments and student attitudes to science in both countries, differences between countries for different scales of the WIHIC were revealed. For example, Indonesian students were more positive about the Involvement and Investigation scales, while Australian students were more positive about the Task Orientation and Equity scales.

The WIHIC has also been used in studies in a number of additional countries, including South Korea, Turkey, the United Arab Emirates (UAE) and South Africa. In South Korea, Kim et al. (2000) used the WIHIC and QTI to assess learning environments and teachers' interpersonal behaviour. They surveyed 543 students in 12 South Korean schools using Korean translations of both questionnaires. The results of the study supported the cross-cultural validity of the Korean version of the WIHIC. Each scale of the WIHIC was found to display satisfactory internal consistency reliability, with Cronbach's alpha values between 0.82 and 0.92. Factor analysis confirmed the a priori factor structure of the WIHIC.

In Turkey, Telli, Cakiroglu, and Brok (2006) used the WIHIC to explore relationships between 1,983 Turkish students' perceptions of the classroom learning environment and their attitudes towards biology. This study found positive relationships between attitudes and the Investigation and Task Orientation scales. It also showed clear differences in attitudes towards biology between girls and boys.

In South Africa, Aldridge, Fraser and Ntuli (2009) used the WIHIC in a study aimed at improving teaching practices among pre-service teachers undertaking a distance education programme in South Africa. A primary version of the WIHIC was developed and translated into IsiZulu. It measured four dimensions: Teacher Support, Involvement, Task Orientation and Equity. Thirty-one teachers administered it to 1,077 students in order to examine their perceptions of their preferred and actual classroom learning environments. This version of the WIHIC showed satisfactory factorial validity and internal consistency reliability.

In the UAE, MacLeod et al. (2010) developed and validated an Arabic translation of the WIHIC. It was administered to 763 college students in 82 classes. This version of the WIHIC showed sound factorial validity and internal consistency reliability for both actual and preferred forms. Also in the UAE, Afari et al. (2013) used a 6-scale, 48-item version of the WIHIC in a study administered to 352 college students in 33 classes (Afari et al., 2013, p. 141), that aimed to determine whether games in college-level mathematics classes improved the students' perceptions of and attitudes towards mathematics. Following principal axis factoring, the number of items was reduced to 46. This study also provided evidence to support the validity and reliability of the WIHIC in a range of settings.

The WIHIC has also been used extensively in the USA, where studies have consistently provided evidence for its validity and the reliability in a variety of settings (den Brok, Fisher, Rickards, & Bull, 2006; Helding & Fraser, 2013; Martin-Dunlop & Fraser, 2007; Ogbuehi & Fraser, 2007; Wolf & Fraser, 2008). It has also been used in a number of common types of learning environment research, including: associations between outcomes and environments (den Brok et al., 2006; Helding & Fraser, 2013; Martin-Dunlop & Fraser, 2007; Ogbuehi & Fraser, 2007); the evaluation of educational innovations (Martin-Dunlop & Fraser, 2007; Wolf & Fraser, 2008); and teacher education (Helding & Fraser, 2013; Martin-Dunlop & Fraser, 2007). Two of these studies (Ogbuehi & Fraser, 2007; Wolf & Fraser, 2008) also replicated past research showing strong and consistent associations between student attitudes and classroom learning environment scales.

This section has described 8 classroom learning environment instruments which have been developed and widely used over the last 50 years. It has focussed on the WIHIC, which is used in this study, providing evidence to support its validity and reliability across a wide range of countries, schools and classrooms.

2.4.3 Past Research Involving Classroom Learning Environment Instruments

Past research involving the use of classroom learning environment instruments provides a wide range of validated instruments and research across at least 10 domains: associations between classroom environment and outcomes; evaluation of educational innovations; differences between students' and teachers' perceptions of classrooms; comparisons of actual and preferred environments; effect on classroom environment of antecedent variables (for example, gender, year level, school type, subject); the transition from primary to secondary school; school psychology; teacher education; educational productivity research; and using environment instruments to facilitate changes in classroom life (Fraser, 2012b). The aims of this research study are to explore relationships between students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment, and to determine whether there are differences in students' perceptions of the classroom learning environment based on their year level, membership of extension or mainstream groups and strand of mathematics being studied. Thus the aspects of past research relevant to the study were environment-outcome associations and determinants of classroom learning environments. The following section presents research relating to these two aspects of past classroom learning environment research:

- associations between the classroom learning environment and student outcomes, which examines studies related to how classroom learning environments affect student outcomes (Section 2.4.3.1), and
- determinants of classroom learning environments, which presents studies interested in finding differences in students' perceptions of the classroom learning environment based on gender, year level, and confidence (Section 2.4.3.2).

2.4.3.1 Associations Between Classroom Learning Environment Perceptions and Student Outcomes Such as Confidence

Analysis of associations between students' cognitive and affective learning outcomes and their perceptions of the psychosocial characteristics of the classroom has been the strongest tradition in classroom learning environment research since its inception. This research has provided predictive validity of students' perceptions in accounting for the amount of variance in learning outcomes beyond that attributable to student characteristics such as general ability (Fraser, 2012a). The value of this type of research is "that attention to creating positive classroom environments is likely to pay off in terms of improving student outcomes" (Fraser, 2019, p. 2).

Extensive lists (Fraser, 1994, 2012b) of classroom learning environment research have been published, providing evidence of the relationships between the classroom learning environment and student outcomes. They have been conducted predominantly in science classes, but also in mathematics, geography and computer education courses, and have included outcomes such as attitude, achievement, satisfaction, academic efficacy, anxiety and enjoyment. The studies have used different instruments, worked with students from different age groups ranging from kindergarten to college, and been undertaken in a number of different countries and languages.

The research which is relevant to this study is that which has shown relationships between classroom learning environments and self-beliefs such as confidence (Kleitman & Gibson, 2011; Yang, 2015), mathematics self-esteem (Chionh & Fraser, 2009; Tran, 2012), self-concept (House, 1975; Trautwein, Lüdtke, Köller, & Baumert, 2006) and academic self-efficacy (Aldridge & Fraser, 2008; Dorman, 2001; Dorman & Adams, 2004; Dorman et al., 2006). The studies of most interest to this research are those showing links between aspects of the classroom learning environment and academic self-efficacy, a self-concept-like construct closest in nature to confidence as defined for this study. Some examples of these studies are reviewed below.

This area of research was largely neglected until Lorsbach and Jinks (1999. p. 157) called for classroom learning environment researchers to explore the relationships between academic efficacy and classroom learning environments, arguing that “student self-efficacy beliefs regarding academic performance can have important implications for improving learning environments and, consequently, student outcomes”. For them, “self-efficacy influences students’ perceptions of the learning environment” and therefore “focusing on students’ academic self-efficacy could alter student perceptions of the learning environment” (Lorsbach & Jinks, 1999, p. 157).

Dorman (2001) was the first to investigate the associations between classroom learning environment and academic efficacy, which he said “must assess students’ perceptions of their competence to do specific activities” (Dorman, 2001, p. 245). The study was conducted in Australia using 1,055 Year 8, 10 and 12 mathematics students in nine schools. Dorman (2001) used the WIHIC and CLES to measure students’ perceptions of the classroom learning environment, plus a seven-item scale from the Patterns of Adaptive Learning Survey (PALS) which included items such as: “I’m certain that I can master the skills taught in maths this year” and “I can do even the hardest work in this maths class if I try”. Simple correlation and multiple regression analyses were carried out using two units of analysis: the student and the year level. There were small, positive, statistically significant associations between academic efficacy and all scales of the WIHIC with the student as the unit of analysis. These were larger when the year group was used as the unit of analysis. Task Orientation, Investigation, Involvement and Teacher Support and Equity accounted for the most variation in Academic Efficacy. Dorman (2001) also found that the WIHIC was a better predictor of Academic Efficacy than the CLES. Other studies with mathematics (Dorman & Adams, 2004) and science (Dorman et al., 2006) classes have replicated these results, leading to the conclusion that “classrooms characterized by high levels of cooperation, harmony, genuine teacher support, student cohesiveness, task orientation and equity are more likely to enhance the confidence of students” (Dorman & Adams, 2004, p. 80).

In Singapore, Chionh and Fraser (2009) used the WIHIC to investigate associations between classroom learning environment and three affective variables – achievement, self-esteem and attitudes – using 2,310 Year 10 mathematics and geography students. They reported that self-esteem and attitudes were higher in classes with more Teacher Support, Task Orientation and Equity.

Aldridge, Afari, et al. (2012) conducted a study in the United Arab Emirates (UAE) involving 352 students in 3 colleges. They used the Teacher Support scale from the WIHIC and the Personal Relevance scale from the CLES, translated into Arabic, and an attitude instrument that measured enjoyment of mathematics classes and academic efficacy. The academic efficacy scale included items similar to those in the instrument used in this study, such as “I find mathematics easy” (I learn mathematics easily) and “I feel I will pass mathematics with ease” (I expect to do well in mathematics). The study reported the existence of a positive relationship between academic efficacy and the two classroom learning environment scales.

Yang (2015) used a modified version of the WIHIC translated into Chinese and two scales selected from the FSMAS to explore the associations between students’ attitudes and their perceptions of the classroom learning environment. The sample consisted of 2,455 students in 52 classrooms from 12 schools in 3 western provinces of China. The two scales used to measure attitudes were Confidence In Learning Mathematics, “which is intended to measure confidence in one's ability to learn and to perform well on mathematical tasks (Fennema & Sherman, 1976, p. 326) and Usefulness Of Mathematics. The FSMAS confidence scale includes items similar to those used in this study, such as “I am sure that I can learn mathematics” (I learn mathematics easily) and “I can get good grades in mathematics” (I expect to do well in mathematics). Yang (2015) reported that Teacher Support and Equality, Involvement and Investigation accounted for a significant amount of the variation in students’ confidence in their ability to do and learn mathematics.

2.4.3.2 Determinants of the Classroom Learning Environment

Classroom learning environment dimensions have been used as criterion variables in research aimed at identifying how a student’s perceptions of their environment might vary. According to Fraser (2002, p. 14) “the determinant of classroom environment that has been most extensively researched in Asia is student gender”. The findings have generally supported the view that female students perceive the classroom learning environment more favourably that male students. Other factors which have also been researched include teacher education level (Helding & Fraser, 2013), grade/year level (D. B. Byrne et al., 1986; Peer, 2011; Peer & Fraser, 2015; Randhawa & Michayluk, 1975), subject (Chionh & Fraser, 2009), school environment (Fraser, 1994) and ability grouping (Costello, 1986; Peer, 2011; Peer & Fraser, 2015; Quek, Wong, & Fraser, 2005a, 2005b; Rita & Martin-Dunlop, 2011). This study is interested in determining whether there are differences in female students’ perceptions of the classroom learning environment based on which year level they are in, and whether they are in an extension or mainstream mathematics class. In order to do this, the scales of the WIHIC were used as dependent variables. This section will review some of the relevant previous research studies.

Differences Between Perceptions of Classroom Learning Environment Based on Year Level

A number of studies have found that students’ perceptions of the classroom learning environment change with age. Yang (2015) found that Year 9 students viewed mathematics class environments less favourably than Year 7 and 8 students. However, Dorman and Fraser (2009) reported that students' perceptions of the classroom learning environment improved for 7 of the 10 scales of the TROFLEI.

Randhawa and Michayluk (1975) used the LEI to conduct a study involving 96 Saskatchewan rural and urban classes. They reported that the differences in mean scores between Grades 8 and 11 were statistically significant for twelve of the fifteen scales of the LEI. They concluded that "grade 8 and 11 classrooms perceive the social environment differently" (Randhawa & Michayluk, 1975, p. 276). They attributed some of the differences to the more formal and rigid approach used in Grade 8 classes when compared with Grade 11 classes. These results are supported by the work of Koul, Fisher, and Shaw (2011) and D. B. Byrne et al. (1986).

D. B. Byrne et al. (1986) studied 1,675 students in 18 schools in New South Wales (NSW), Australia. They used the MCI, CES and ICEQ to measure the students' preferred classroom learning environments and the Quality of School Life (QSL) to measure their perceptions of the actual environment. The sample consisted of students in Years 7, 9 and 11. They found some interesting gender and year level differences. They reported that the girls placed more emphasis on social harmony and less on competitiveness than did the boys. In terms of year levels, Year 7s favoured structure and cohesiveness, while Year 9s preferred friction and competition and Year 11 students more independence. The results show that rather than a decline in perceptions of the classroom learning environment as they progress through school, there is a change in student preference for types of environment, from one that is structured to one that provides more independence.

Studies of students' transition from primary school to secondary school have also found differences in students' perceptions of the classroom learning environment based on year level (Deieso, 2016; Feldlaufer, Midgley, & Eccles, 1988; Hine, 2001; Midgley et al., 1989; Telli et al., 2006). Hine (2001) and Deieso (2016) reported that students in South Australian schools had less positive perceptions of the classroom learning environment following their transition to high school. P. Ferguson and Fraser (1999) found that students viewed the high school classroom learning environment more favourably than the primary classroom learning environment in that there was less friction and competitiveness, but also that teacher–student relationships were perceived to have deteriorated. Feldlaufer et al. (1988) reported that students had fewer opportunities to cooperate among themselves and to be actively involved in the learning process.

A study of 1,081 primary school science students (Grades 4, 5 and 6) in Singapore (Peer & Fraser, 2015) found statistically significant differences of small magnitude for gender, grade level and stream, where stream refers to the Gifted Education (GE) stream versus the High Ability (HA) stream. The data were collected using 10 scales of the WIHIC and the Test of Science Related Attitudes (TOSRA). The authors reported statistically significant differences between grade levels for Teacher Support, Task Orientation and Cooperation. Generally Grade 6 students had a more favourable perception of the classroom learning environment than their younger school mates.

The results reported above suggest that students' perceptions of the classroom learning environment differ over time. But rather than deteriorating or improving over time, those perceptions change as students mature, with older students preferring classrooms with more independence and younger students preferring classrooms with more structure.

Differences Between Perceptions of Classroom Learning Environment Based on Ability Grouping (Extension vs Mainstream)

Quek et al. (2005a, 2005b) used the Chemistry Learning Environment Instrument (CLEI) and the QTI to investigate differences between gifted and non-gifted students' perceptions of teacher–student interactions. They surveyed 497 (200 gifted and 297 non-gifted) tenth grade chemistry students in 3 independent schools in Singapore. They reported that gifted students generally regarded their chemistry learning environment more favourably and had a more favourable view of the teachers than their non-gifted counterparts. The differences were statistically significant for Student Cohesiveness, Integration and Material Environment from the CLEI and for Student Responsibility/Freedom and Strict Behaviour from the QTI.

Rita and Martin-Dunlop (2011) conducted a study in California with 146 gifted and 115 non-gifted tenth grade biology students. They used seven scales of the WIHIC to collect data on the students' perceptions of their actual and preferred classroom learning environments. Using the students' perceptions of the actual classroom learning environment, they found that the means for the non-gifted students were lower for all seven scales. The differences were statistically significant for six of the seven scales, with Student Cohesiveness being the odd one out. Teacher Support had the greatest impact on how the students perceived their actual classroom learning environments.

Peer (2011) and Peer and Fraser (2015) studied 1,081 primary school students in 55 classes at 4 schools in Singapore, and reported that students in the Gifted Education group perceived the classroom learning environment differently from students in the High Ability group, with statistically significant differences for the Involvement, Cooperation and Personal Relevance sales.

Fraser and Lee (2009) and S. Lee, Fraser, and Fisher (2003) studied 439 South Korean students in 13 classes that were divided into 3 ability groups based on science ability/preference. They reported that the students in the high ability group perceived the classroom learning environment more favourably than those in the middle and low groups.

This section has traced the history of research into classroom learning environments from its beginnings in the 1920s to the present day. It has described the development and use of eight questionnaires for use in classroom learning environment research, and in particular the WIHIC, which was chosen for use in this study. A modified version of the WIHIC was used because of its ability to elicit information about students' perceptions of their personal involvement in a class. It was important to obtain data about their personal perceptions of the classroom learning environment in order to explore its relationship with their confidence in their ability to do and learn mathematics. Various studies have been identified which attest to the validity and reliability of the WIHIC in a variety of settings, countries and languages. Two of the types of learning environment research identified by Fraser (2012b) are relevant to this study: associations between student outcomes and classroom learning environment;

and criterion validity of classroom learning environment perceptions (e.g., year level and ability level). Studies relevant to each type have been identified and show that there are relationships between classroom learning environments and aspects of self such as academic self-efficacy, year level and ability group.

2.5 Summary

This chapter has reviewed some of the extensive body of research relevant to this study, which examines the existence of relationships between girls' perceptions of their classroom learning environments and their confidence in their ability to do and learn mathematics. Section 2.2 examines the gender gap with regard to girls' lower levels of participation in mathematics and other STEM subjects, while Section 2.3 explores some of the issues relating to the use of different self-belief terms, in particular self-concept, self-confidence (confidence) and self-efficacy, as well as how confidence can be measured and its relationship to achievement, year level, ability group and strand of mathematics studied. A detailed review of previous and current research related to classroom learning environments is provided in Section 2.4.

Section 2.2 provides a review of literature relating to the gender gap – the differences in achievement and participation rates observed between girls and boys. This is relevant to my study, as I have focussed on girls' confidence in their ability to do and learn mathematics in the hopes of identifying aspects of the classroom learning environment that may be related it. The literature reviewed provides evidence to support the view that while the gender gap in achievement has narrowed, the gender gap in participation rates still exists today.

Section 2.3 reviews literature pertaining to confidence and some of the problems associated with the use of this and other self-belief terms. Section 2.3.1 examines definitions of confidence and their use in some of the early studies of girls and mathematics (Fennema & Sherman, 1976, 1977; Reyes, 1984; Reyes & Stanic, 1988; Tapia & Marsh, 2004). It examines the relationship between confidence and other self-beliefs, in particular self-concept and self-efficacy, by dividing self-beliefs into two categories – self-concept-like beliefs and self-efficacy-like beliefs Marsh et al., 2019) – and proposing that confidence be defined as a self-concept-like construct which refers to self-beliefs about one's abilities to do and learn mathematics. This is global in that it accounts for students' beliefs in their overall performance in a particular strand of mathematics, rather than their ability to answer specific mathematics questions. This section concludes by outlining the development of the ATMI and reviews studies which have used it. Section 2.3.2 considers the relationship between confidence and achievement, year level, ability grouping and strand of mathematics studied. The general consensus is that a positive relationship exists between achievement and confidence. Lack of consistency in research findings can be attributed to the variations in terminology used, the instruments used to measure confidence and the different ways of measuring achievement. The literature reviewed suggests that there is evidence to show that confidence declines with age. A large part of the literature relating to ability groupings relates to tracking or setting, and the suggestion is that while confidence appears to be greater for those in higher ability classes, this may in fact be a result of the student being placed in the that group. There was little literature available relating to confidence and strand of mathematics studied.

Section 2.4 reports on literature relating to classroom learning environment research. Section 2.4.1 reviews the background to the field of classroom learning environments, from the early work of Lewin (1936) and Murray (1938) through to the work of Walberg and Anderson (1968b), who developed the first instrument to measure students' perceptions of classroom learning environments. It then progresses into the 1990s, which saw the development of many more such instruments. Section 2.4.2 reviews the development and validation of a number of classroom learning environment instruments including, to name just a few, the LEI, the CES and the COLES. Section 2.4.2.8 provides a detailed review of the WIHIC, from which six scales were chosen for this study. Several studies supporting the validity and reliability of the WIHIC with a wide range of students in different countries are described. Studies which explore similar relationships to this study are discussed in Section 2.4.3, which examines two of the many different types of research in this area: associations between student outcomes and their perceptions of the classroom learning environment in section 2.4.3.1, and determinants of classroom learning environments in section 2.4.3.2.

The next chapter describes the research methodology used in this thesis.

Chapter 3

METHOD

3.1 Introduction

While Chapter 2 provided an overview of the literature related to this study, this chapter outlines the research design and methods used. The chapter is organised into the following sections:

- research questions (Section 3.2);
- research methods and design (Section 3.3);
- the sample (Section 3.4);
- instruments used for data collection (Section 3.5);
- data analysis (Section 3.6);
- ethical considerations (Section 3.8); and
- chapter summary (Section 3.9).

3.2 Research Questions

The research questions guiding this study were introduced in Chapter 1. They are:

1. Are the surveys to be used to collect data for the present study valid and reliable for use in girls' Catholic schools in metropolitan South Australia?
2. Do relationships exist between students' confidence in their ability to do and learn mathematics and their
    a. perceptions of their classroom learning environment?
    b. achievement scores in mathematics?
3. Do students in different year levels differ in terms of their
    a. perceptions of their classroom learning environment?
    b. confidence in their ability to do and learn mathematics?
4. Do students in extension classes differ from other students in the same year level in terms of their
    a. perceptions of their classroom learning environment?

b. confidence in their ability to do and learn mathematics?

5. Do confidence levels differ for different strands of mathematics?

3.3 Research Methods

The overarching focus of the research reported in this thesis is to explore the relationships that might exist between the variables ‘confidence and classroom learning environment’ and ‘confidence and achievement scores’ (Question 2), and the differences between students’ perceptions of the classroom learning environment and their confidence in their ability to do and learn mathematics based on membership of a particular group (Questions 3 to 5). In his list of the major characteristics of quantative research, Creswell (2012, p. 13) includes the comparison of groups or relating variables using statitsical analysis. Therefore, following the advice of L. Cohen et al. (2013, p. 604) that the use of quantitative data analysis “is entirely dependent on fitness for purpose”, this study employs a quantitative research method.

Surveys were selected as an appropriate means of collecting data because they provide an efficient and economical way of collecting large amounts of data from a wide range of people at the same time. As such, they provided a means to

> gather data at a particular point in time with the intention of describing the nature of existing conditions, or identifying standards against which existing conditions can be compared, or determining the relationships that exist between specific events. (L. Cohen et al., 2013, p. 256)

In this study, data were gathered from female middle school students about their perceptions of their classroom learning environments and their confidence in their ability to do and learn mathematics. The data were then analysed using a variety of statistical techniques to explore whether there were any relationships between the students’ perceptions of the classroom learning environment and their confidence.

Collecting data about perceptual measures is a useful way of determining what students think and feel about the classroom learning environment. Describing the classroom learning environment in terms of students' perceptions "has the dual advantage of characterising the setting through the eyes of the participants themselves and capturing data which the observer could miss or consider unimportant" (Fraser, 1998b, p. 8). Students have considerable experience of classroom learning environments, having spent approximately 20,000 in classrooms throughout their school careers (Fraser, 2001).

Using a correlational survey design, four steps were taken in order to collect the data:

- The selection and modification of surveys.
- A pilot study to check the face value of the surveys following the changes made during the selection process.
- Three rounds of data collection. In Round One data were collected from 335 students. These data were used in answering Research Questions 1 to 5. In Rounds Two and Three data were collected from approximately 200 students from School One. These data were used to answer Research Question 5. Each round of data collection was undertaken at the completion of a unit of work, and students were asked to reflect on their classroom experiences while that unit was being taught. Generally, a unit of work is concerned with only one strand of mathematics.
- Recording of an achievement score for each unit. At the end of each unit of work students were assessed using a common test written by their teachers. Their grades for these tests were recorded as their achievement scores for each unit.

3.4 The Sample

This section describes the sample and how it was selected, including the selection of study sites (Section 3.4.1). The composition of the sample is described in Section 3.4.2 and the pilot study is described in Section 3.4.3.

3.4.1 Site Selection

Six Catholic girls' schools in Adelaide were approached to participate in the study; four responded to the invitation, two of which agreed to participate. I chose girls' schools because confidence is frequently given as a reason for girls and women being underrepresented in STEM-related subjects and subsequently occupations (Eccles, 2011b; Hart, 1989; Watt, 2007), and I wanted to eliminate the presence of boys in the class as a variable (see Chapter 1 for more information).

Both schools were situated in the eastern suburbs of Adelaide, and were within a kilometre of each other. The first school, referred to as School One, was a Reception to Year 12 boarding school with approximately 800 students on 2 campuses. Students in the middle school (Years 6 to 10) participated in this study. At the time of this study two mathematics courses were provided: mathematics and extended mathematics, both of which used the International Baccalaureate-Middle Years Programme (IB-MYP) mathematics curricula. These curricula identify five strands of mathematics: Number, Algebra, Geometry and Trigonometry, Statistics and Probability, and Discrete Mathematics. In School One, the Extended Mathematics programme was offered to approximately 15% to 20% of students at each year level.

Students were selected for extension classes based on their prior achievement in mathematics. In Years 6 to 9, selected students were withdrawn from their normal mathematics class once a week for extension activities based on the Extended Mathematics curriculum. In Year 10 there was a separate class using the IB-MYP Extended Mathematics curriculum that was intended to prepare students for the study of two units of mathematics in Year 12. All other students were randomly assigned to classes. Two hundred and three middle school students (Years 6 to 10) in 15 classes participated in the study.

School Two was a Year 8 to Year 12 day school with approximately 350 students. This school used the South Australian Curriculum Standards and Accountability (SACSA) framework and the Australian Curriculum to determine subject content. The Australian Curriculum provides explicit descriptions of what is to be taught (content) and what is expected in terms of the quality of learning by year level or band of schooling (achievement standards), from Foundation to Year 12. (ACARA, 2010). In School Two students were randomly assigned to mathematics classes, with a small group of students at Year 10 provided with a less rigorous course. One hundred and thirty-four middle school students in 10 classes participated in the study.

The sample therefore comprised 335 middle school students in 25 classes from 2 Catholic girls' schools in metropolitan Adelaide. The next section describes the make-up of the sample.

3.4.2 Participants

All middle school students in both schools (Years 6 to 10 in School One and Years 8 to 10 in School Two) were invited to participate in the study. Despite difficulties in getting students to return consent forms, the final sample was 335. This sample was drawn from 25 classes (15 from School One and 10 from School Two). The sample represented about 60% of all students in the relevant year levels.

A breakdown of the sample included: 11 Year 6 students from 2 classes, 44 Year 7 students from 2 classes, 81 Year 8 students from 6 classes, 111 Year 9 students from 7 classes and 100 Year 10 students from 8 classes.

Table 3.1 Sample profile

Year level	Number of students in each year level	Number of students in extension class	Number of classes
6	10	3	2
7	45	13	2
8	81	15	6
9	102	18	7
10	97	18	8

3.4.3 Pilot Sample

Before the surveys were administered to the main sample, a pilot study was undertaken in order to check the face validity of the items and ensure that the meanings had been retained despite changes to some items in the surveys used (see Sections 3.5.1.1 and 3.5.1.2 for information related to changes to the surveys). The sample for the pilot study involved one Year 9 mathematics class of 18 students. This class was chosen because the teacher was willing to participate and because the researcher was known to some of the students, making it more likely that they would respond honestly to the surveys. On the day selected for the trial, the 10 students who were at school completed the surveys and then participated in an informal discussion of the items with the researcher. The results of the discussion are provided in Section 3.5.2.

3.5 Instruments Used for Data Collection

In this section details are provided about the collection of the data. Two surveys were used: one to assess students' confidence in their ability to do and learn mathematics (described in Section 3.5.1.1), and one to assess students' perceptions of the classroom learning environment (described in Section 3.5.1.2). The collection of achievement data for each student is described in Section 3.5.1.3. Details of the pilot study used to examine the face validity of individual items are described in Section 3.5.2, and finally, the administration of the two surveys is described in Section 3.5.3.

3.5.1 The Instruments

This section describes the instruments used to collect data for the present study. Scales and items from the ATMI and the Student Adaptive Learning Engagement Scale (SALES) were used to collect data relating to students' confidence in their ability to do and learn mathematics. Section 3.5.1.1 describes the selection process and the modifications. To collect data relating to students' perceptions of the classroom learning environment scales from two surveys, the WIHIC and the COLES were used. This process is described in Section 3.5.1.2.

3.5.1.1 The Attitude Towards Mathematics Inventory (ATMI)

Confidence is defined in this study as a self-concept-like construct that refers to self-beliefs about one's ability to do and learn mathematics. It is global in that it accounts for students' beliefs in their overall performance in a particular strand of mathematics, rather their ability to answer specific mathematics questions. In choosing an instrument to measure confidence, those measuring attitudes to mathematics with a confidence scale were considered. This included one of the most popular and widely used (Pepin, 2011), the FSMAS, which was developed in 1976. However, this instrument was considered to be rather dated and too long, taking about 45 minutes to administer. It was thought that this would cause significant disruption to classes and teaching programmes. It was also feared that many of the students involved in multiple data collection sessions would withdraw due to the time needed to complete the survey. The more recently developed, and shorter, ATMI was chosen. The self-confidence scale of the ATMI was designed to assess the confidence and self-concept of students with respect to their ability to do and learn mathematics (Tapia & Marsh, 2004). It includes items such as "I have a lot of self-confidence when it comes to mathematics", "I learn mathematics easily" and "I am confident that I could learn advanced mathematics". These items fit well with the definition of confidence being used in this study.

The other attitudes measured by the ATMI are Value, Enjoyment and Motivation, where Value measures students' beliefs about the worth of mathematics in their daily lives now and in the future, Enjoyment measures how much students enjoy doing mathematics, and Motivation measures students' interest in mathematics and their likelihood to study it in the future (Tapia & Marsh, 2004, p. 17). The scales have different numbers of items: Confidence has 15, Enjoyment and Value have 10 and Motivation has 5, making a total of 40 items. These items are responded to using a five-point Likert scale – a "multiple-indicator or multiple-item measure of a set of attitudes relating to a particular area" (Bryman, 2012, p. 166) – of Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree.

As outlined in Chapter 2, the ATMI was developed "to investigate the underlying dimensions of attitudes toward mathematics" (Tapia & Marsh, 2004, p. 17) "that may contribute to math anxiety and to expand beyond measurement of enjoyment" (Tapia & Marsh, 2005, p. 273). The original development, which involved 545 high school students, showed it to be reliable and valid for that sample. Other studies using the ATMI have confirmed its reliability and validity with more samples, including: Tapia and Marsh (2005), who used it with 134 college aged American students; Lim and Chapman (2013), who developed a short form of it for a predominantly Chinese sample in order to provide an updated and shorter instrument for use in Singapore; Asante (2012), who used it with 181 Ghanaian students when exploring differences between boys' and girls' attitudes to mathematics; Majeed et al. (2013), who conducted a confirmatory factor analysis of it with a sample of 699 students in South Australia; and Khine and Afari (2014), who explored its reliability and validity with 269 middle-school students in the UAE.

After review, the ATMI's four scales (Confidence, Value, Enjoyment and Motivation) were maintained. However, the number of items in each scale was reduced to six, with only positive items being retained. Negatively worded items are sometimes used to avoid passive responses and bias (Velayutham et al., 2011); however, there is evidence (Schriesheim, Eisenbach, & Hill, 1991) that positively worded items improve internal consistency and response accuracy. This process reduced the total number of items from 40 to 23. Some items were then reworded or replaced with similar items from other instruments to make them more student-friendly.

The changes to the Confidence and Value scales included reducing the number of items to six per scale. In choosing which items to retain, only positive statements were considered. These were reworded as necessary to use language more in tune with the students in the sample. For example, "I can think of many ways that I use mathematics outside of school" became "I can think of ways to use mathematics in my daily life". The meanings of these two scales remained the same, with Confidence measuring the students' confidence and self-concept with respect to their ability to do and learn mathematics and Value measuring their beliefs about the worth of mathematics in their daily life now and in the future.

Enjoyment measures how much students enjoy doing mathematics. The wording of many of the items was similar to that of the Attitude to Subject scale of the COLES (Bell & Aldridge, 2014). Six of the seven items from this scale were chosen and reworded to include the word mathematics. "I really like mathematics" became "I look forward to mathematics lessons", and "I am happier in a math class than in any class" became "Mathematics is one of my favourite school subjects".

In the ATMI, Motivation measures the students' interest in mathematics and their likelihood of studying it in the future. As students' future intentions were not relevant to this study, however, I decided to use items from the Learning Goal orientation scale of the SALES instead, as this measures the students' interest in mathematics. The SALES was developed by Velayutham et al. (2011) to measure factors related to motivation and self-regulation in science learning in middle-school students. The name of the Motivation scale was retained and the items were re-worded for mathematics. There is a full copy of the m-ATMI in Appendix A. Table 3.2 (below) provides a description of the m-ATMI used in this study.

Table 3.2 Description and scale for each m-ATMI scale

Scale Name	Scale Description	Sample Item
Confidence	… the students' confidence and self-concept of their performance in mathematics.	I am confident when it comes to mathematics.
Value	… the students' beliefs about the worth of mathematics in daily life now and in the future.	Mathematics is a worthwhile subject.
Enjoyment	… how much students enjoy doing mathematics.	I look forward to mathematics lessons.
Motivation	… the students' interest in mathematics.	I want to learn new mathematics content.

Each statement is responded to using a five-point Likert scale with the alternatives: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree.

3.5.1.2 Students' Perceptions of the Classroom Learning Environment

This section describes the selection of the scales used to determine students' perceptions of the classroom learning environment. Six of the seven scales used were selected from the WIHIC. The WIHIC was developed by Fraser et al. (1996) to provide an instrument that would be better suited to assessing differences in perceptions of the classroom learning environment held by different students or groups of students in the same class. The WIHIC included both scales that had previously been shown to be significant predictors of outcomes (Fraser, 1994) and new ones that addressed the then recent changes in classroom teaching. The first version of the WIHIC included 90 items in 9 scales. Further refinement of the WIHIC followed with Aldridge et al.'s (1999) study involving 1,081 students from Western Australian schools and 1,879 students from Taiwanese schools. This version of the WIHIC, which has 54 items in 7 scales, was used as the basis for the survey in this study.

The WIHIC is "worded to elicit the student's perception of his/her individual role within the classroom, as opposed to the student's perception of the class as a whole" (Dorman, 2008, p. 181). This was important to this study, as I wanted to explore the relationship between each student's perception of the classroom learning environment and their confidence in their ability to do and learn mathematics.

Another reason for using the WIHIC in this study was that its reliability and validity in its various forms have been replicated in numerous studies involving different countries, languages and age groups, including: 1,081 Australian students using an English version and 1,897 Taiwanese students using a Chinese version (Aldridge et al., 1999); 763 students in UAE using an Arabic version with primary-aged students (MacLeod & Fraser, 2010); 1,077 primary-aged students in South Africa (Aldridge et al., 2009); 978 secondary students in Queensland, Australia (Dorman, 2008); 250 adults in Singapore (Khoo & Fraser, 2008); 2,043 Year 11 and 12 students in Western Australia (Aldridge, Fraser, et al., 2012); and many other studies as listed by Fraser (2012b, p. 1206). This survey was considered to be a suitable choice because "the majority of these studies also furnished evidence of the ability of the WIHIC to differentiate between the perceptions of students in different classrooms" (Fraser, 2012b, p. 1206).

Close scrutiny of the WIHIC was undertaken to ensure that the scales were all relevant to the present study, the language was appropriate for the range of students in the sample (Years 6 to 10) and the administration of the two surveys at the same time would not take too long. I was looking for items that would reflect the teacher's ability to create the right environment for students to find mathematics more interesting and enjoyable (Fraser & Fisher, 1982; Goodenow, 1993) and to give students greater confidence in their ability to succeed in the classroom (Goodenow, 1993). "Children learn better when they are actively engaged" (Stones, 1972, p. 392) in the learning process, and "Students are more likely to do well in their learning if they feel accepted and do not experience harassment and prejudice from either the teacher or their peers" (Aldridge, Fraser, et al., 2012, p. 267): seventeen years of teaching girls had led me to believe that a number of factors within the classroom learning environment might be related to students' confidence in their ability to do and learn mathematics, and these include the relationship between the student and the teacher, the relationship between the student and the rest of the class and the degree to which the student took responsibility for their own learning.

With this in mind, six of the seven scales from the WIHIC were chosen for this study. The Investigation scale, which assesses the degree to which students engage in the skills and processes of enquiry, was deemed to be unsuitable as there were few investigations carried out in the mathematics classes included in the sample. Of the chosen WIHIC scales, Teacher Support and Equity speak to the relationship between the teacher and the student, Student Cohesiveness and Cooperation provide insights into the student's perception of the relationships between the members of the class, and Involvement and Task Orientation focus on the students' participation in their own learning.

One aspect of student participation in learning which is not included in the WIHIC is assessment. Burton (1992, p. 7) believed that the three dimensions of syllabus, pedagogy and assessment "work together in close interaction to define learning". A 2012 study involving action research, which focussed on teacher development, also highlighted the role of students in their learning. The COLES was developed for use in the 2012 study (Aldridge, Fraser, et al., 2012). It included two scales that focused on assessment. One of these scales, Clarity of Assessment Criteria, was chosen for this study. It measures the extent to which assessment criteria are made clear to the students. "I know how to complete assessment tasks successfully" is an example of one of the items in the scale. Aldridge, Fraser, et al. (2012) concluded that assessment becomes educative when students recognise the link between it and their learning. Assessment then contributes positively to their learning and their ability to plan for future learning. It also encourages self-directed learning practices.

When the seven scales had been decided on, the wording and number of items in each scale was also reviewed. Two scales, Student Cohesiveness and Task Orientation, were re-named Student Friendship and Task Focus respectively to make them more meaningful to the students. The items pertaining to each scale were grouped together. A five-point frequency response scale (Almost Never, Seldom, Sometimes, Often, Almost Always) was used. Eventually 6 items per scale were retained, making a total of 42 items reduced from 54. Table 3.3 (below) provides details of the WIHIC used to collect data in this study. A full copy of the WIHIC is provided in Appendix B.

Table 3.3 Description and sample for each WIHIC scale

Scale name	Scale description	Sample item
	The extent to which …	
Student Friendship/ Cohesiveness	students support and help each other.	I make friendships among students in this class.
Teacher Support	the teacher supports the students.	The teacher helps me when I have trouble with my work.
Equity	students are treated fairly by the teacher.	I get the same amount of help from the teacher as do other students.
Involvement	students participate in the lessons and what other work they might do.	I explain my ideas to other students.
Task Focus /Orientation	students think it is important to complete set work and to stay on task.	Getting a certain amount of work done is important to me.
Cooperation	students cooperate rather than compete with each other in the class.	We work in groups (or pairs) in this class.
Clarity of Assessment Criteria	assessment criteria are made clear to the students.	I know what types of information are needed to complete an assessment task.

Each statement is responded to using a five-point frequency scale with the alternatives: Almost Never, Seldom, Sometimes, Often, and Almost Always.

3.5.1.3 Achievement Data

Achievement scores were provided by the teachers at the end of each unit of work. The data collected were the scores obtained by each students on the common topic test, written by the teachers, that they took at the end of a unit of work. A note was made of the strand of mathematics to which the unit belonged. These data were used to answer Research Question 2.

3.5.2 Pilot Study

As both the WIHIC and the ATMI had been modified to make them suitable for the study described in this thesis, it was necessary to undertake a pilot study prior to administering the surveys to the whole sample. The aims of the pilot study were: to examine the face validity of the individual items; to determine whether the proposed sample was capable of completing the surveys; to identify whether there were items in the survey that were likely to make the respondents feel uncomfortable; to check the time taken to complete the survey; and to provide an opportunity for the researcher to gain some experience in administering the surveys, as recommended by Bryman (2012) and Creswell (2012).

On the morning of the pilot study the students were sent an email which contained the link to the surveys in SurveyMonkey. The students had been asked not to complete the surveys outside of their mathematics class. Once in the class, the students were asked to open the email and click on the link. They were immediately connected to the instructions for the WIHIC survey. Once they had completed the first survey (WIHIC), they were given instructions on how to complete the second survey (ATMI). I read the instructions to the students to ensure that they were thinking about their mathematics classes when answering the survey questions. A copy of the instructions for both surveys is included in Appendix C.

Once the instructions had been given, students completed the surveys in their own time. They were allowed to ask the supervisor for clarification if needed but not allowed to speak to each other. Once all of the students had completed the online surveys, I held an informal discussion with them in order to obtain their reactions to the surveys as described in the previous paragraph.

Face validity addresses concerns about “whether an indicator appears to reflect the content of the concept in question” (Bryman, 2012, p. 711), or in other words, whether the respondents have the same understanding of the questions as the researcher. Students were asked to explain what they understood by particular questions. The students were able to explain their understanding of the items and why they had chosen their responses, which suggested that they understood the questions as intended. Other comments made by the students in response to questions about the surveys included “straightforward; easy to answer the questions; well set out”. It was clear from their responses that this group of students was capable of completing the surveys. There was no evidence that any of the questions made them feel uncomfortable. However, there was a request to explain the term ‘Seldom’, one of the responses to the WIHIC survey. In talking with the students, it was confirmed that that they did not have any problems with understanding the questions.

The time for each student to complete the surveys was recorded in SurveyMonkey. Completion times ranged from 7 minutes to 12 minutes, with the average time taken being 10.10 minutes. This was considered an appropriate length that would not be too disruptive to timetabled lessons, as students would only miss about half of a 45-minute lesson each time they completed the surveys (allowing for time to get to and from the computer laboratory).

The final purpose of the pilot study was to provide an opportunity for the researcher to identify any problems with the administration process. The main problem identified was that the students found the instructions too long, and preferred to receive verbal instructions from the supervisor. This was useful information that was acted on when administering the surveys to the whole sample.

3.5.3 Administering the Surveys

The two surveys were administered at the same time, and for ease of administration and subsequent data entry it was decided to use an electronic survey delivered using the platform SurveyMonkey. The use of an electronic survey had a number of advantages, including reduced paper usage and a means of ensuring that students to answer all questions (as it was not possible to move on until a question had been answered). Using SurveyMonkey's Email Invitation Collector, which is designed to help track respondents by including their email addresses and IP addresses in the survey results, also provided additional advantages. These included: an increased sense of confidentiality (as students were not required to enter any personal details); data being available almost as soon students had completed the surveys; a reduction in data entry errors, as data was available electronically; and the ability for students to easily change answers without the resulting confusion caused by crossed-out answers etc. It also facilitated data matching once the achievement data had been acquired.

The final survey as entered into SurveyMonkey was organised into three parts. Part 1 asked students about their mathematics class membership, and the mathematics topic they were currently studying or had just finished studying; this data was used to answer Research Question 5. Part 2 asked the students about the classroom learning environment using the selected scales and items from the WIHIC. All items for each scale appeared on the same page. The students were asked to complete the questions using Actual and Preferred options and answered using the frequency format of Almost Never, Seldom, Sometimes, Often and Almost Always to indicate how often the particular behaviour occurred in their class. These appeared together, as shown in Figure 3.1 below.

*1.	I make friendships among students in this class				
	Almost never	Seldom	Sometimes	Often	Almost Always
Actual	○	○	○	○	○
Preferred	○	○	○	○	○

Figure 3.1 Response format of the WIHIC

Once the students had completed Part 2 there was another set of instructions on how to complete the final section. Part 3 asked the students questions about their attitudes towards mathematics using the selected scales and items of the m-ATMI. For Part 3, all items for each scale appeared on the same screen. Answers were given via a five-point Likert Scale using the terms Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree. An example of the questions in Part 3 is given in Figure 3.2 below.

*46.	I have a lot of self-confidence when it comes to mathematics.				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	○	○	○	○	○

Figure 3.2 Response format of the m-ATMI

The survey was disseminated to each student using SurveyMonkey's Email Invitation Collector. Each student was thus provided with a unique survey link on the same day that they were to complete the survey, which facilitated the linking of responses for those students who completed the questionnaire multiple times and for adding the achievement data. Despite students receiving an email invitation, they were asked not to complete the survey in their own time.

The survey was administered online during mathematics lessons and was supervised by either the researcher or the mathematics teachers. The decision to supervise the completion of the surveys was made for two reasons. The first was to ensure that students were answering the questions in relation to mathematics lessons, and the second was to avoid the low response rates that often result from unsupervised online surveys. A number of studies have highlighted the problems associated with the completion of online surveys. Dogan (2012) suggests participation can be as low as 22% of those invited to participate. A study by Bray and Maxwell (1985) in which a number of paper-based supervised surveys were compared with unsupervised online surveys showed that participation was between 8% and 36% lower for the studies involving unsupervised online surveys. Dillman (2006) suggests that five contacts with participants are necessary to get a reasonable response rate. Students who were absent when their class completed the survey were encouraged through email reminders to complete it online; very few took this opportunity, which supported the above concerns with regard to unsupervised completion of the surveys. About 25% of absentees subsequently completed the surveys. In order to provide data to answer Research Question 5, two further rounds of data were collected in a similar way.

Achievement data – the results of a common test taken at the completion of the unit of work – were provided by the teachers. These were matched to the relevant students via email addresses.

3.6 Data Analysis

This section describes the methods used to analyse the data and find answers to the research questions. The variables to be considered in the data analysis were confidence, year level, ability group, strand of mathematics being studied, the scales measured by the surveys and achievement scores. Year level, ability group, strand and achievement score were categorical variables, while m-ATMI and WIHIC scales were continuous variables.

3.6.1 Reliability and Validity of the m-ATMI and the WIHIC

To ensure that each survey was valid and reliable for this South Australian sample, the data were analysed to examine factorial validity, scale validity (alpha reliability coefficient) and the ability to differentiate between classes (analysis of variance (ANOVA)). For both surveys, principal axis factoring with oblique rotation was used to examine the internal structure. This involved actual and preferred versions of the 7-scale WIHIC. The criteria for retention of an item were that it had a factor loading of .40 on its *a priori* scale and less than .40 on all other scales in the same survey. The individual student was used as the unit of analysis, and items not meeting the criteria above were removed. The factors were rotated "to foster their interpretability" (T. A. Brown, 2014, p. 27), with low correlations suggesting little overlap with other scales.

To examine the degree to which items in the same scale measure the same aspect, the reliability of each scale of the m-ATMI and WIHIC was estimated for two units of analysis: individual and class means. The internal consistency reliability of each scale was calculated separately for the m-ATMI and the actual and preferred data of the WIHIC using Cronbach's alpha coefficient. A reliability coefficient of 0.7 is considered satisfactory, while 0.8 is good (Pallant, 2013).

An analysis of variance (ANOVA) "is used when you have two or more groups and you wish to compare their mean scores on a continuous variable" (Pallant, 2013, p. 190) It compares the variance between groups – in this case the year levels – for each classroom learning environment scale for the actual version of the WIHIC. Conducting an ANOVA enabled an examination of whether each WIHIC scale could differentiate between the perceptions of students in different classrooms. The process calculates an F ratio, which is the variance between groups divided by the variance within groups. A large F ratio implies that there is more variability between groups than within groups.

3.6.2 Relationship Between Confidence, Classroom Learning Environment and Achievement

To answer Research Question 2, simple correlation and multiple regression analyses were used. The simple correlation provided information about bivariate associations between students' confidence in their ability to do and learn mathematics and each of the classroom learning environment scales. Multiple regression, R, allows a more sophisticated exploration of the relationships, providing information about which variables are the best predictors of outcomes. R^2 tells "how much of the variance in the dependent variable is explained by the model" (Pallant, 2013, p. 166). The *beta* value is used to compare the contribution of each independent variable.

The simple and multiple regression analyses were performed separately using the individual student as the unit of analysis. To examine the relationship between students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment, the calculations used the four scales of the m-ATMI as the dependent variables and the six scales of the WIHIC as the independent variables. To examine the relationship between students' confidence in their ability to do and learn mathematics and their achievement score, simple correlation analyses and multiple regression analysis were also used, with Confidence as the independent variable and the achievement score as the dependent variable.

3.6.3 Differences Between Confidence and Perceptions of the Classroom Learning Environment based on Year/Grade Level or Ability Level

This section outlines the data analysis used to answer Research Questions 3 and 4. To examine whether differences exist for students based on group membership in terms of their perception of the classroom learning environment and their confidence in their ability to do and learn mathematics, one-way multivariate analysis of variance (MANOVA) was used. MANOVA is used when there is more than one dependent variable. In this case there were several, as the scales of the m-ATMI and the WIHIC were the dependent variables. Year level and ability group (extension or mainstream) were used as the independent variables. MANOVA was used to compare the responses of students in different groups and to determine whether the differences between the groups were statistically significant. It is preferable to conducting a series of ANOVAs separately, as this could inflate the Type 1 error. The MANOVA produced an *F*-value which compared variation among the groups with variation within the groups. The *F*-values were considered to be significant if the corresponding significance value generated at the same time was less than alpha divided by the number of dependent variables to be investigated.

The importance of the impact of year level or ability group on the classroom learning environment and Confidence can be evaluated using the effect size statistic. The effect size or strength of association can then be calculated using the formula for Cohen's d, which is:

$$d = (m_1 - m_2)/\sqrt{(s_1^2 + s_2^2)/2},$$

where m_1 and m_2 are the means and s_1 and s_2 are the standard deviations of Group 1 and Group 2. Cohen's d gives the differences between the groups in terms of the standard deviation units. J. Cohen (2013, p. 25) proposed that the effect size was small for d = .2, medium for d = .5 and large for d = .8

If the F ratio indicates that there are significant differences among the groups, post-hoc tests can be performed to determine where these differences occur. Post-hoc tests are designed to guard against increasing the Type 1 error. In this study, Tukey's Honestly Significant Difference (HSD) test was used to determine between which groups the differences occurred. For Research Question 3 it was performed between year levels, and for Research Question 4 it was performed between extension and mainstream classes.

Before performing MANOVA, assumption testing was conducted to check for normality, linearity, homogeneity of variance-covariance matrices and multicollinearity. To test multicollinearity, correlations between the dependent variables were found. These ranged from .20 between Student Friendship/Cohesiveness and Task Focus/Orientation to .63 between Task Focus/Orientation and Clarity of Assessment Criteria for the scales of the WIHIC and from .49 between the Value and Confidence scales to .77 between the Value and

Motivation scales for the m-ATMI, suggesting the moderate correlations required for MANOVA to work best. A few violations were noted in the tests for normality. The sample is negatively skewed, with student responses tending to be in the upper range of possible answers for both the m-ATMI and the WIHIC. The Mahalanobis distance, which is the distance of a particular case from the centroid of remaining cases, identified 5 cases out of 335 as multivariate outliers. Further examination of these cases showed that they were not that different from the remaining cases and that the 5% trimmed mean and the mean values were very similar, and so these cases were retained.

3.6.4 Differences in Confidence Based on Strand of Mathematics Being Studied

MANOVA was also used in order to explore differences in students' confidence in their ability to do and learn mathematics for different strands of mathematics. In this study, a group of students completed the survey on three occasions at the end of three different units of work. Three groups were identified as having completed the same three strands. They were analysed separately. The independent variable was the strand of mathematics being studied and the dependent variables were the scales of the m-ATMI. The technique provides information about significant differences between the three sets of scores, if they exist.

3.7 Ethical Issues

"Ethical issues arise in survey research at distinct points in the research process, such as in collecting data, in analysing results and in reporting results" (Creswell, 2012, p. 402). To ensure that all ethical issues involved in this research were appropriately handled, it was necessary to apply to the university's ethics committee for approval. The study was deemed to be low-risk research in line with the *National Statement on Ethical Conduct in Human Research 2007* (Updated May 2013), which defines low-risk research as "research in which the only foreseeable risk is one of discomfort" (NHMRC, 2013, p. 16). The approval document is included in Appendix D. This section outlines the ethical issues considered when designing this research project. These include: harm to participants and informed consent in Section 3.7.1; privacy and confidentiality in Section 3.7.2; and disruption to the site in Section 3.7.3 (Bryman, 2012).

3.7.1 Harm to Participants and Informed Consent

Stress could be the main cause of harm to participants. For the students, stress could be caused by disruption to their normal routines or by concerns about the confidentiality of their answers: would the teacher know what they had said about the classroom learning environment? For the teachers, stress might be caused by the increased scrutiny that their teaching would face. The differences between their perceptions of the classroom learning environments they had created and those of the students could contribute to their stress and possible loss of self-esteem. They could also be worried about what information would be shared with the principal and other members of the staff. Another concern for teachers could be whether there would be extra work for them and the amount of disruption there might be to their teaching programmes.

Informed consent implies that the choice to participate is freely made and that there are no negative consequences for non-participation. The choice to participate can only be freely made if the participant has sufficient information about the study and its possible effects on them. Consent was obtained from the principals to gain access to their schools, the teachers whose classes were involved in the study, the students and their parents.

In order to reduce harm to the participants and to ensure informed consent, information sheets were provided to the principals, teachers, parents and students. Initially a letter (Appendix E) was written to the principal of each school requesting permission to conduct the research in their school. This letter was accompanied by copies of the Student and Teacher Information Sheets and the survey (Appendices H & I). Follow-up phone calls were made as necessary. Two schools agreed to participate in the study and the principals directed me to work with the mathematics coordinators. The information sheets contained all information about the data to be collected and how it would be collected, and the measures that would be taken to protect confidentiality and maintain the privacy of the participants. Students were given the choice to participate or not, and could withdraw from the study at any time. They were told that there would be no penalty for those who chose not to participate or who withdrew at a later stage. Students were given information sheets and consent forms (Appendix H) to take home. Students and parents were required to sign the consent forms. Provision was made for parents to contact me or the school if they had further questions or required any clarification. Consent forms were returned to mathematics teachers and then passed on to me. Teachers were also provided with a similar information sheet (Appendix I). I also attended a mathematics faculty meeting with the teachers to explain the purpose of the research, provide all the relevant information and allay any fears teachers may have had. I was at pains to point out that the only extra work for them would be collecting the consent forms and providing student achievement scores at appropriate times. I also promised to minimise disruption to their lessons, which I did by being punctual and being prepared for data collection in those cases where I supervised the students.

3.7.2 Privacy and Confidentiality

Issues of privacy are linked to confidentiality and informed consent in that the participant will expect whatever information they provide to be confidential, i.e., kept private. In this study the main issues were about who would see the data, how privacy would be protected and how the data would be stored and used. As mentioned above, the information sheets pointed out that the only people who would see the data would be me and my supervisor. All data is stored on my password-protected laptop. Consent forms have been stored at Curtin University, where they will remain for seven years, after which they will be destroyed. Only summaries of the students' responses were provided to the teachers at School One.

3.7.3 Disruption to the Site

Long periods of disruption to lessons can contribute to both student and teacher stress. Disruption can take a number of forms: students missing all or part of a lesson; increased pressure on school resources (in this case access to computers); uncertainty about when data will be collected; which room to go to for data collection. Disruption to the normal timetable and programmes was kept to a minimum. Students undertook the surveys during mathematics lessons. They were supervised either by me or by their mathematics teachers. In School 2, where the teachers supervised the administration of the survey, I provided a set of instructions for the supervisors (Appendix J). In both schools disruption to normal lessons was kept to a minimum.

3.8 Summary

This chapter describes and justifies the research methods and design employed in this study. It also includes descriptions of the sample, the instruments used for data collection, the data analysis and ethical issues related to the study.

The study sits within the positivist paradigm and relies solely on the collection and analysis of quantitative data. The design included a small pilot study to check the face value of the surveys following changes to the chosen instruments. Three rounds of data collection ensued. A sample of 335 students in 25 classes drawn from Year 6 to 10 students attending two girls' Catholic colleges in Adelaide, South Australia was surveyed in Round 1, and a smaller sample of 200 in 15 classes from one school for Rounds 2 & 3. The data from Round 1 were used to answer Research Questions 1 to 4, with combined datasets from Rounds 1, 2 and 3 being used to answer Research Question 5.

The data were collected using surveys based on the ATMI and the WIHIC. The ATMI was used to determine their attitudes towards mathematics, in particular their confidence in their ability to do and learn mathematics, and the WIHIC was used to assess the students' perceptions of their classroom learning environments. Modifications were made to each instrument to make them more appropriate for the sample. These changes are described in detail, together with the reasons for them. The m-ATMI included 24 items, with 6 items in each of the 4 scales Confidence, Value, Enjoyment and Motivation. For convenience, the surveys were administered at the same time.. The WIHIC included 42 items with 6 items in each of 7 scales Student Friendship/Cohesiveness, Teacher Support, Equity, Involvement, Task Focus/Orientation, Cooperation and Clarity of Assessment Criteria. The m-ATMI was responded to using a five-point Likert scale with the terms Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree. The response scale for the WIHIC involved a five-point frequency response of Almost Never, Seldom, Sometimes, Often, Almost Always and Always. The surveys were delivered using SurveyMonkey.

The chapter further discusses the ethical considerations addressed in undertaking this study. The first consideration was that no harm should be done to the participants by causing them stress. The possible causes of stress were considered, and all attempts were made to alleviate them. This was primarily done by providing detailed information about the study and the impact it would have on the participants to school principals, teachers, parents and students. The provision of this information also ensured that teachers and students could give informed consent to participate in the study. Participants were informed that participation was voluntary and that they could withdraw at any time without any negative consequences. They were also advised that no aspect of the study would be used in determining their grades in mathematics. All students and teachers were assured that no-one would see their responses or class data except for me and my supervisor. Disruption to the site was also considered, and every effort was made to minimise this by liaising closely with relevant mathematics and other staff.

The next chapter reports the data analyses and findings of my study.

Chapter 4

DATA ANALYSIS

4.1 Introduction

The purpose of this chapter is to describe the data analysis and findings with respect to the five research questions introduced in Chapter 1. The results of the analysis of the data are reported in the following sections. Section 4.2 presents evidence to support the reliability and validity of the scales used to assess students' self-reporting of self-confidence and their perceptions of the classroom learning environment. Section 4.3 reports the results of the analysis used to examine the relationships between students' confidence in their ability to do and learn mathematics, their perceptions of the classroom learning environment and their achievement. The next three sections report the results of analyses that examine differences in students' confidence and their perceptions of the classroom learning environment based on membership of particular groups: year level in Section 4.4; ability group in Section 4.5; and strand of mathematics being studied in Section 4.6. Finally, a chapter summary is provided in Section 4.7.

4.2 Reliability and Validity of the Instruments

As a first step, it was important to ensure that the instruments used to collect the data were reliable and valid when used with the sample selected for this study. Therefore, Research Question 1 asked:

> Are the surveys to be used to collect data for the present study reliable and valid for use in girls' Catholic schools in metropolitan South Australia?

Prior to analysis, the suitability of data was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity. According to Pallant (2013, p. 193), the KMO value should be above .60 and "Bartlett's test of sphericity should be statistically significant at $p < .05$". For these data, the KMO score was .90, which is well above the .60 recommended, and Bartlett's Test of Sphericity was 0.00, which also complies.

Data collected from 335 students in 25 classes at two Catholic girls' schools in Adelaide, South Australia were used to provide evidence to support the reliability of the surveys used in the present study. The data analysis methods used to check for reliability and validity were described in Chapter 3. This section reports the evidence for both the m-ATMI (Section 4.2.1) and the WIHIC (Section 4.2.2).

4.2.1 Reliability and Validity of the m-ATMI

This section reports the results of the analyses carried to determine the reliability and validity of the m-ATMI. The first section (Section 4.2.2.1) examines the factor structure of the m-ATMI. The second section (Section 4.2.2.2) reports on the internal consistency reliability (Cronbach's alpha) of the m-ATMI. The final section (Section 4.2.2.3) reports on the strength of the correlations between the factors of the m-ATMI, which provides information on the discriminant validity of the m-ATMI.

4.2.1.1 Factor Structure of the m-ATMI

Principal axis factoring with oblique rotation was used to examine the internal structure of the 23-item, 4-scale m-ATMI as described in Chapter 3. In accord with the recommendations of T. A. Brown (2014) and Tabachnick and Fidell (2013), no items were problematic, as all had factor loadings of .40 or more on their own scales and less than .40 on other scales. Factor loadings for the four scales of Confidence, Value, Enjoyment and Motivation are reported in Table 4.1 below. The bottom of the table reports the percentage of variance and the eigenvalues for each scale. The percentage of variance ranged from 4.06% to 52.86% and explained 72.72% of the total variance. As per Kaiser's criterion, the eigenvalue for each factor is greater than 1, and hence all scales are retained.

Table 4.1 Factor loadings, eigenvalues, and percentage of variance for the m-ATMI scales

Item No	Factor loading			
	Confidence	Value	Enjoyment	Motivation
1	.87			
2	.81			
3	.91			
4	.78			
5	.78			
6	.91			
7		.54		
8		.72		
9		.66		
10		.93		
11		.83		
12		.67		
13			.84	
14			.93	
15			.73	
16			.71	
17			.82	
18			.77	
19				.42
20				.79
21				.75
22				.81
23				.66
Eigenvalue	2.70	12.16	1.26	1.01
%variance	11.74	52.86	4.06	4.06

Loadings less than 0.40 have been omitted
N = 335 students in 25 classes

4.2.1.2 Internal Consistency Reliability of the m-ATMI

The internal consistency reliability indicates whether the items in the scale assess a similar construct. For this study, Cronbach's alpha coefficient for each scale was calculated to provide an estimation of the internal consistency reliability. Table 4.2 reports the Cronbach's alpha coefficient for each of the four scales of the m-ATMI with the individual and class means as the units of analysis. All the coefficients were above .85, which is well above the .70 recommended as the appropriate level (Pallant, 2013). This confirms that the survey has high internal consistency reliability for this sample.

Table 4.2 Internal consistency reliability (Cronbach's alpha coefficient) for the m-ATMI with the individual and class means as the units of analysis

Scale	No. of items	Unit of analysis	Cronbach's alpha
Confidence	6	Individual	.93
		Class mean	.92
Value	6	Individual	.92
		Class mean	.96
Enjoyment	6	Individual	.94
		Class mean	.91
Motivation	5	Individual	.89
		Class mean	.85

N = 335 students in 25 classes

4.2.1.3 Discriminant Validity of the m-ATMI Scales

The discriminant validity of a scale refers to the degree to which it measures what it says it measures. It is determined by examining the correlations of related and unrelated variables. The strength of the correlations between the factors of the m-ATMI is given in Table 4.3. According to T. A. Brown (2014, p. 116), "a factor correlation that exceeds .80 or .85 is often used as the criterion to define poor discriminant validity". The values ranged from .36 to .62, all of which are below .80.

Table 4.3 Component correlation matrix for the m-ATMI scales

	Value	Confidence	Enjoyment	Motivation
Value	–			
Confidence	.42	–		
Enjoyment	.54	.62	–	
Motivation	.61	.36	-.52	–

N = 335 students in 25 classes
Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

4.2.2 Reliability and Validity of the WIHIC

This section reports the results of four analyses carried out to determine the reliability and validity of the WIHIC. Section 4.2.2.1 examines the factor structure of the WIHIC. Section 4.2.2.2 reports on the internal consistency reliability (Cronbach's alpha) of the

WIHIC. Section 4.2.2.3 reports the results of a one-way analysis of variance which provides information about the ability of the WIHIC to differentiate between classes. Section 4.2.2.4 reports on the strength of the correlations between the factors of the WIHIC, which provides information on the discriminant validity of the WIHIC.

4.2.2.1 Factor Structure of the WIHIC

Principal axis factoring with oblique rotation was used to examine the internal structure of the 42-item, 7-scale version of the WIHIC. Oblique rotation "provides a more realistic representation of how factors are interrelated" than orthogonal rotation. It also "yields a more accurate representation of the magnitude of these relationships" (T. A. Brown, 2014, p. 28). The goal of principal axis analysis is to explain as much of the variance in variables as possible using as few components as possible. Factor loadings represent the strength of the correlation between variables and factors. Tabachnick and Fidell (2013) recommend that in order to be statistically meaningful, factor loadings should be at least .32. T. A. Brown (2014, p. 27) says that factor loadings between .3 and .4 "are interpreted as salient"; in other words, the variables are meaningfully related to the factor. For an item to be retained it had to have a loading of at least .40 on its *a priori* scale and less than .40 on each of the other six WIHIC scales for both actual and preferred versions. The factor loadings for the actual and preferred versions of the WIHIC (analysed separately) are reported in Table 4.4 below. During item analysis, the six items in the Equity scale and seven other items were found to be problematic and were removed from further analysis. The seven items removed were: items 3 and 5 from the Student Friendship/Cohesiveness scale; item 12 from the Teacher Support scale; item 21 from the Involvement scale; item 30 from the Task Focus/Orientation scale; item 36 from the Cooperation scale; and item 37 from the Clarity of Assessment Criteria scale. As a result, 29 items in 6 of the 7 modified WIHIC scales were retained. The results reported in Table 4.4 show that with one exception, the remaining items, for both the preferred and actual versions all loaded at .40 or more for their own scales and below .40 on all other scales. The exception was Item 8. For the preferred version it loaded at .30 on its own scale and at .42 on the Task Focus/Orientation scale. However, it was retained because its removal weakened the overall factor structure. The results provide strong support for the factorial validity of the instrument when used with this sample.

The bottom of Table 4.4 reports the percentage of variance for each scale in the actual and preferred versions of the WIHIC and their eigenvalues, which "represent the amount of the total variance explained by that factor" (Pallant, 2013, p. 191). According to Kaiser's criterion only factors with an eigenvalue greater than 1 should be retained. For both the actual and preferred versions of the remaining six scales of the WIHIC, all of the retained factors have eigenvalues greater than 1. These factors explain 59.94% and 59.43% of the variance for each version respectively.

Table 4.4 Factor loadings, eigenvalues, and percentage of variance for WIHIC scales

Item No.	Factor loadings											
	Student Friendship/ Cohesiveness		Teacher Support		Involvement		Task Focus/ Orientation		Cooperation		Clarity of Assessment Criteria	
	Act	Pref	Act	Pref	Act	Pref	Act	Pref	Act	Pref	Act	Pref
1	.69	.57										
2	.68	.73										
4	.75	.74										
6	.63	.46										
7			.70	.61								
8			.77	.30				.42				
9			.73	.80								
10			.71	.62								
11			.75	.78								
19					.75	.68						
20					.83	.81						
22					.69	.82						
23					.68	.65						
24					.68	.72						
25							.75	.79				
26							.64	.73				
27							.73	.71				
28							.73	.73				
29							.66	.70				
31									.65	.81		
32									.65	.50		
33									.76	.89		
34									.65	.49		
35									.53	.52		
38											.71	.73
39											.62	.80
40											.81	.83
41											.52	.60
42											.81	.82
Eigen-values	1.44	1.37	1.96	1.53	1.18	8.40	2.54	1.86	1.73	1.71	8.53	2.36
% of variance	5.00	4.74	6.75	5.27	4.08	29.00	8.74	6.43	5.98	5.90	29.43	8.15

N = 335 students in 25 classes
Loadings less than .40 have been omitted.

4.2.2.2 Internal Consistency Reliability of the WIHIC

As with the m-ATMI, the Cronbach's alpha reliability coefficient was used an as index of internal scale consistency. The results reported in Table 4.5 indicated that for the actual version of the WIHIC, the coefficients ranged from .70 to .87 with the individual as the unit of analysis and from .71 to .92 with the class mean as the unit of analysis. For the preferred version, the values ranged from .63 to .85 with the individual as the unit of analysis and .74 to .93 with class mean as unit of analysis. Apart from the coefficient for the individual mean of the preferred scale for Student Friendship/Cohesiveness, all the coefficients were above .70, which suggests that the survey has high internal consistency reliability for this sample.

Table 4.5 Internal consistency reliability (Cronbach's Alpha coefficient) for the WIHIC using the individual and class means as the units of analysis

Scale	No of Items	Unit of Analysis	Alpha Reliability	
			Actual	Preferred
Student Friendship/Cohesiveness	4	Individual	.70	.63
		Class Mean	.71	.74
Teacher Support	5	Individual	.85	.79
		Class Mean	.87	.74
Involvement	5	Individual	.84	.84
		Class Mean	.92	.93
Task Focus/Orientation	5	Individual	.78	.83
		Class Mean	.78	.86
Cooperation	5	Individual	.74	.74
		Class Mean	.76	.83
Clarity of Assessment Criteria	5	Individual	.87	.85
		Class Mean	.91	.87

N = 335students in 25 classes

4.2.2.3 Ability of WIHIC to Differentiate Between Classes.

To determine whether the scales of the WIHIC were able to differentiate between the perceptions of students in different classes, a one-way analysis of variance (ANOVA) was performed using the data related to the actual version of the WIHIC. The independent variable was class membership, with the 335 students divided between 25 classes. Table 4.6 (below) reports the eta^2 statistic, which is the ratio of “between” to “total” sums of squares and represents the proportion of variance in scale scores accounted for by class membership, and provides an estimate of the strength of the association between the independent variable (class membership) and the dependent variable (WIHIC scale). The eta^2 statistic ranged from .07 to .20 for different WIHIC scales. Statistically significant differences ($p < .01$) between students’ perceptions in different classes were reported for four of the six scales: Teacher Support, Involvement, Cooperation and Clarity of Assessment Criteria. This suggests that these four scales were able to differentiate between students’ perceptions in different classes. The variance due to class membership for these variables ranged from 11% (Clarity of Assessment Criteria) to 20% (Teacher Support).

Table 4.6 Ability of the Actual Version to Differentiate Between Classes (ANOVA Results) for the WIHIC

Scale	ANOVA results (eta2)
Student Friendship/Cohesiveness	.07
Teacher Support	.20**
Involvement	.15**
Task Focus/Orientation	.07
Cooperation	.16**
Clarity of Assessment Criteria	.11**

N = 335 students in 25 classes
***p* < .01

4.2.2.4 Discriminant Validity for the WIHIC

Table 4.7 shows the strength of the correlations between the factors of the WIHIC for the actual and preferred versions. The values to the right of the diagonal are the results for the actual version and the values to the left of the diagonal are the results for the preferred version. According to T. A. Brown (2014, p. 116), “a factor correlation that exceeds .80 or .85 is often used as the criterion to define poor discriminant validity". For these data, the values range from .10 to .41 for the actual version and .20 to .35 for the preferred version, which are well below .80, suggesting good discriminant validity for the WIHIC.

Table 4.7 Component correlation matrix for the WIHIC

	Student Friendship/ Cohesiveness	Teacher Support	Involvement	Task Focus/ Orientation	Cooperation	Clarity of Assessment Criteria
Student Friendship/ Cohesiveness	----	.14	.30	.10	.22	.15
Teacher Support	.25	----	.30	.30	.18	.30
Involvement	.26	.35	----	.25	.27	.40
Task Focus/ Orientation	.24	.27	.30	----	.18	.41
Cooperation	.21	.27	.29	.23	----	.25
Clarity of Assessment Criteria	.20	.23	.29	.34	.22	----

4.3 Relationships Between Confidence, Classroom Learning Environment and Achievement

One of the main objectives of this study was to explore whether relationships exist between students' confidence in their ability to do and learn mathematics, their perception of the learning environment and their achievement as measured by the topic tests given by each school. The data collected using the m-ATMI and the WIHIC were analysed to address Research Question 2:

> Do relationships exist between students' confidence in their ability to do and learn mathematics and their
>
> a. perceptions of their classroom learning environment?
> b. achievement grade in mathematics?

The following sections report the results of the analysis carried out to answer the questions above. Simple correlation and multiple regression analyses were performed on the data collected from students and teachers. Section 4.3.1 describes the relationships between the students' confidence in their ability to do and learn mathematics as measured by the m-ATMI and their perceptions of the classroom learning environment as measured by the WIHIC. The results are reported in Table 4.8. Section 4.3.2 reports the results of the multiple regression analysis between the scales of the m-ATMI and the students' achievement as measured by the school-based tests. These results are recorded in Table 4.9.

4.3.1 Relationships Between Students' Confidence in Their Ability to Do and Learn Mathematics and Their Perceptions of Their Learning Environment

This section describes the results of the analysis performed to explore the relationship between the students' confidence in their ability to do and learn mathematics and their perceptions of the learning environment. Two techniques, simple correlation and multiple regression, were used on the data collected from 335 students in 25 classes. Correlation is a statistical technique used to describe the strength and direction of the linear relationship between two variables and is given by r. Multiple regression is based on correlation but allows for a more sophisticated exploration of relationships that may exist in the data. It can also tell which variables are the best predictors of outcomes.

The level of statistical significance indicates how much confidence we can have in the results obtained for this sample, while the r-values help to determine the strength of the association between the variables. The results of the simple correlation were statistically significant ($p < .01$) for all but two pairings, the exceptions being that of Student Friendship/Cohesiveness and Enjoyment ($p < .05$) and that of Student Friendship/Cohesiveness and Motivation, which were not statistically significant.

J. Cohen (2013, p. 79) suggests that the correlation is small for r-values between .10 and .29, medium for r-values between .30 and .49 and large for r-values greater than .50. Using this as a guide, the results reported in Table 4.8 below support that there was no correlation between Motivation and Student Friendship/Cohesiveness. There was a small correlation between Confidence and Student Friendship/Cohesiveness (r =

.17), Teacher Support (r = .29) and Cooperation (r = .22), and between Value and Student Friendship/Cohesiveness (r = .16) and Enjoyment and Student Friendship/Cohesiveness (r = .13). There was a medium correlation between Confidence and Task Focus/Orientation (r = .38), Value and Teacher Support (r = .32), Involvement (r = .31), Task Focus/Orientation (r = .45), Cooperation (r = .31) and Clarity of Assessment (r = .44), and between Enjoyment and Teacher Support (r = .46), Involvement (r = .41) and Task Focus/Orientation (r = .48), Motivation and Teacher Support (r = .30) and Clarity of Assessment (r = .45). Finally, there was a large correlation between Confidence and Involvement (r = .51) and Clarity of Assessment (r = .56), Enjoyment and Clarity of Assessment (r = .55), and Motivation and Task Focus/Orientation (r = .50). The strongest correlations involve Clarity of Assessment Criteria, Involvement and Task Focus/Orientation.

The β-values (standardised coefficients) can be used to determine which of the variables in the model contribute to explaining the dependent variable. The variable with the largest beta-coefficient makes "the strongest unique contribution to explaining the dependent variable" (Pallant, 2013, p. 167). In this case the β-values showed that the variables making a statistically significant contribution to the prediction of Confidence were Involvement and Clarity of Assessment Criteria. The results in Table 4.8 indicate that the Clarity of Assessment Criteria scale made the greatest unique contribution to the prediction of Confidence, with a value of .43 ($p < .01$). Involvement, with a β-value of .36, made a smaller but still statistically significant contribution to Confidence ($p < .01$). The β-values tell us that if the score for the Clarity of Assessment Criteria scale were to be improved by 1 standard deviation, then Confidence could be improved by 0.43 standard deviations.

The last row in Table 4.8 gives the Multiple Correlation R. For Confidence this was .63. R^2 tells us how much of the variance in Confidence was due to the WIHIC model. In this case, R^2 = .40, which means that 40% of the variance in Confidence was due to the learning environment model.

Table 4.8 Simple correlation and multiple regression analyses for associations between WIHIC and m-ATMI Scales with student as the unit of analysis

Scale	Confidence		Value		Enjoyment		Motivation	
	r	β	r	β	r	β	r	β
Student Friendship/ Cohesiveness	.17**	- .07	.16**	-.05	.13*	-.12*	.09	-.09
Teacher Support	.29**	- .06	.32**	.05	.46**	.18**	.30**	.05
Involvement	.51**	.36**	.31**	.07	.41**	.16**	.27**	.04
Task Focus/Orientation	.38**	.06	.45**	.27**	.48**	.20**	.50**	.36**
Cooperation	.22**	- .02	.31**	.15**	.27**	.05	.23**	.07
Clarity of Assessment Criteria	.56**	.43**	.44**	.16**	.55**	.26**	.45**	.20**
Multiple Correlation (R)		.63**		.52**		.62**		.54**

*$p < .05$ **$p < .01$
N = 335 students in 25 classes

4.3.2 Relationships Between Students' Confidence in Their Ability to Do and Learn Mathematics and Achievement

This section describes the results of the analysis performed to explore the relationship between students' self-reports of their confidence in their ability to do and learn mathematics and their achievement. As in the previous section, two techniques, simple correlation and multiple regression, were used. Teachers provided the achievement data based on topic tests undertaken by students in their classes. The teachers of two classes chose not to participate in this aspect of the study, reducing the sample for this analysis to 308 students in 23 classes.

The results of the simple correlations recorded in Table 4.9 show that the *r*-values were statistically significant ($p < .01$) for the relationships between the achievement score and all four of the m-ATMI scales. The values ranged from .56 (for the relationship between students' self-reports of Confidence and their achievement scores, which had the strongest relationship) to .25 for the relationship between Value and Achievement. Using Cohen's guidelines, reported in Section 4.3.2, there is a large correlation between achievement and Confidence and a medium correlation between achievement and Enjoyment and Motivation, with a small correlation between achievement and value.

The β-value, which was used to determine which variables in the model contributed to the dependent variable, was only statistically significant for the relationship between Confidence and achievement score. This score indicated that Confidence made the strongest unique contribution to the prediction of the dependent variable achievement score. The results also indicated that if Confidence could be increased by 1 standard deviation then the achievement score could be increased by 0.53 standard deviations.

The multiple correlation (R) for achievement score, reported in the last row of Table 4.9, was .57 and was statistically significant ($p < .01$). R^2 tells us how much of the variance in achievement was due to the m-ATMI model. In this case $R^2 = .32$, which means that 32% of the variance in achievement score was due to the m-ATMI model, with Confidence making the greatest contribution.

Table 4.9 Simple correlation and multiple regression analyses for associations between the m-ATMI scales and achievement scores

Scale	Achievement Score	
	r	β
Confidence	.56**	.54**
Value	.26**	-.12
Enjoyment	.41**	.05
Motivation	.32**	.10
Multiple Correlation (*R*)		.57**

**p* < .05 ***p* < .01 (2-tailed)
N = 308 students in 23 classes

4.4 Differences in Student Confidence in Their Ability to Do and Learn Mathematics and Their Perceptions of the Classroom Learning Environment Based on Year Level

The third research objective of this study was to explore whether there were differences in students' confidence in their ability to do and learn mathematics and their perceptions of the learning environment due to their year level. The data collected using the WIHIC and the m-ATMI were analysed using MANOVA to address Research Question 3:

> Do students in different year levels differ in terms of their
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their learning environment ?

"MANOVA tests whether mean differences among groups on a combination of DVs are likely to have occurred by chance" (Tabachnick & Fidell, 2013, p. 245). In this case, the MANOVA was conducted using the actual scales of the modified WIHIC and the scales of the m-ATMI as the dependent variables.

Section 4.4.1 reports the differences between year levels in terms of student confidence in their ability to do and learn mathematics. Due to the small number of Year 6 students participating, they were excluded from this analysis, reducing the sample to 325 students in 23 classes. Section 4.4.2 reports the findings for students' perceptions of learning environments for students in different year levels.

4.4.1 Differences Between Students' Confidence in Their Ability to Do and Learn Mathematics in Different Year Levels

This section reports the differences in students' attitudes to mathematics, and in particular their confidence in their ability to do and learn mathematics, for different year levels. The average means as reported in the first four columns of Table 4.10 were used to portray the differences between year levels for each of the scales of the m-ATMI. Figure 4.1 provides a graphical profile of the average item means for each scale of the m-ATMI for each year level. From inspection of the graphs, it appears that there are differences between the year levels, with Year 7 having higher value means than Year 10. There is very little difference between the standard deviations for each year level and scale.

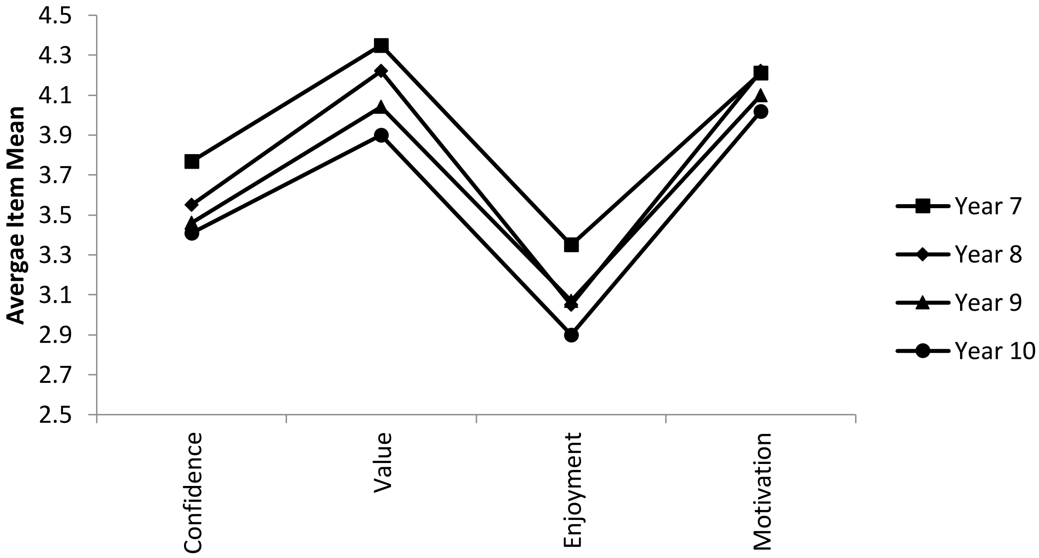

Figure 4.1 Year level means for each scale of the m-ATMI

MANOVA was used to examine whether the differences in item means were statistically significant. The four scales of the m-ATMI (Confidence, Value, Enjoyment and Motivation) were used as the dependent variables, with year level as the independent variable. The results of the MANOVA produced a Wilks' Lambda value of 0.933, with a significance value $p < .05$, which suggests there is a statistically significant effect for year level. This suggests that there is a change in attitude towards mathematics for different year levels.

The final column of Table 4.10 gives the results of the MANOVA: the F value. The F value can be used to indicate whether there are statistically significant differences between the class means for each scale of the m-ATMI. In this case it shows that there are statistical significances between year levels for one scale – Value ($p < .01$) – but not for Confidence, which is the scale of most interest to this study.

Tukey's HSD multiple comparisons procedure was carried out to ascertain whether there were statistically significant differences between particular pairs of year levels. The results are summarised in Table 4.11, in which asterisks indicate pairs of year levels for which there are statistically significant differences. The results indicate that there are statistically significant differences between some year levels for two scales. For the Value scale, with $F = 4.37$ and $p < .01$, there are statistically significant differences between Years 7 and 10 ($p < .05$, effect size = 0.58) and Years 8 and 10 ($p < .05$, effect size = 0.40). For the Enjoyment scale, with $F = 2.36$, there is a statistically significant difference between Years 7 and 10 students ($p < .05$, effect size = 0.48).

Table 4.10 Average item mean, average item standard deviation and MANOVA results for the differences between Years 7, 8, 9 and 10 students attitudes towards mathematics using the m-ATMI

Scale	Average Item Mean				Average Item Standard Deviation				F
	Year 7	Year 8	Year 9	Year 10	Year 7	Year 8	Year 9	Year 10	
Confidence	3.77	3.55	3.46	3.41	1.00	0.99	0.88	0.90	1.66
Value	4.35	4.22	4.04	3.90	0.84	0.87	0.80	0.72	4.37**
Enjoyment	3.35	3.05	3.07	2.90	1.01	0.95	0.95	0.87	2.43
Motivation	4.21	4.22	4.10	4.02	0.78	0.72	0.75	0.69	1.41

*$p < .05$ **$p < .01$
N = 45 students in Year 7, 81 students in Year 8, 102 students in Year 9, 97 students in Year 10.

The magnitudes of the differences between each pair of year levels for the Value and Enjoyment scales of the m-ATMI, which show statistically significant differences between year levels, range from small to medium as per the guidelines for eta^2 specified by J. Cohen (2013, pp. 25–26). These guidelines for group comparisons are: small = .2, medium = .5 and large = .8. Figure 4.2 shows that for the Value scale, the effect size between Years 7 and 10 is just above medium, and small to medium between Years 8 and 10.

Table 4.11 Effect size and Tukey's HSD multiple comparison for statistical significance of difference between each pair of year levels for the m-ATMI

Scale	Effect Size[a] & Statistical Significance					
	Years 7 – 8	Years 7 – 9	Years 7 – 10	Years 8 – 9	Years 8 – 10	Years 9 – 10
Confidence	0.22	0.33	0.39	0.10	0.15	0.06
Value	0.15	0.38	0.58*	0.22	0.40*	0.18
Enjoyment	0.30	0.29	0.48*	-0.02	0.16	0.19
Motivation	-0.01	0.14	0.26	0.16	0.28	0.11

*$p < .05$ **$p < .01$
N = 45 students in Year 7, 81 students in Year 8, 102 students in Year 9, 97 students in Year 10.
[a] the effect size is the difference between two means divided by the pooled standard deviation, providing the magnitude of the difference in terms of standard deviation units.

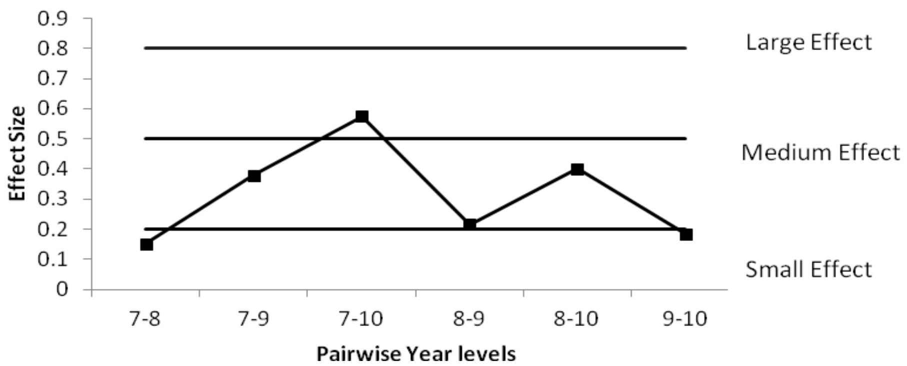

Figure 4.2 Effect size for the Value scale

Figure 4.3 shows that for the Enjoyment scale, the effect size between Years 7 and 10 is medium.

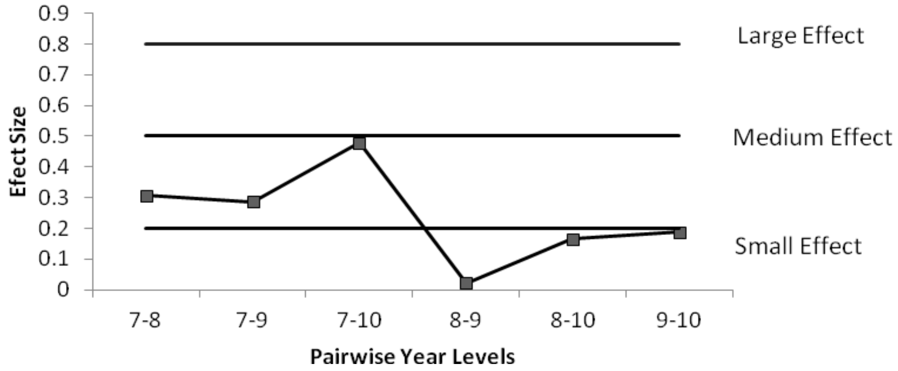

Figure 4.3 Effect size for Enjoyment scale

4.4.2 Differences Between Students' Perceptions of Their Learning Environments in Different Year Levels

This section reports the differences in students' perceptions of the learning environment for different year levels. As a first step, the average item means, reported in the first four columns in Table 4.12, were used to portray the differences between year levels (see Figure 4.4). Table 4.12 below shows the mean and standard deviations for each year level for each of the WIHIC scales. Inspection of the graph suggests that there are differences in means between year levels, in particular between Years 7 and 10. The graph also shows a decrease in mean values for each scale across the year levels. Inspection of the standard deviations for the different year levels indicates that they do not vary considerably, although the standard deviations for Year 10 students' responses are slightly higher. The graph does not provide information about whether these differences are statistically significant.

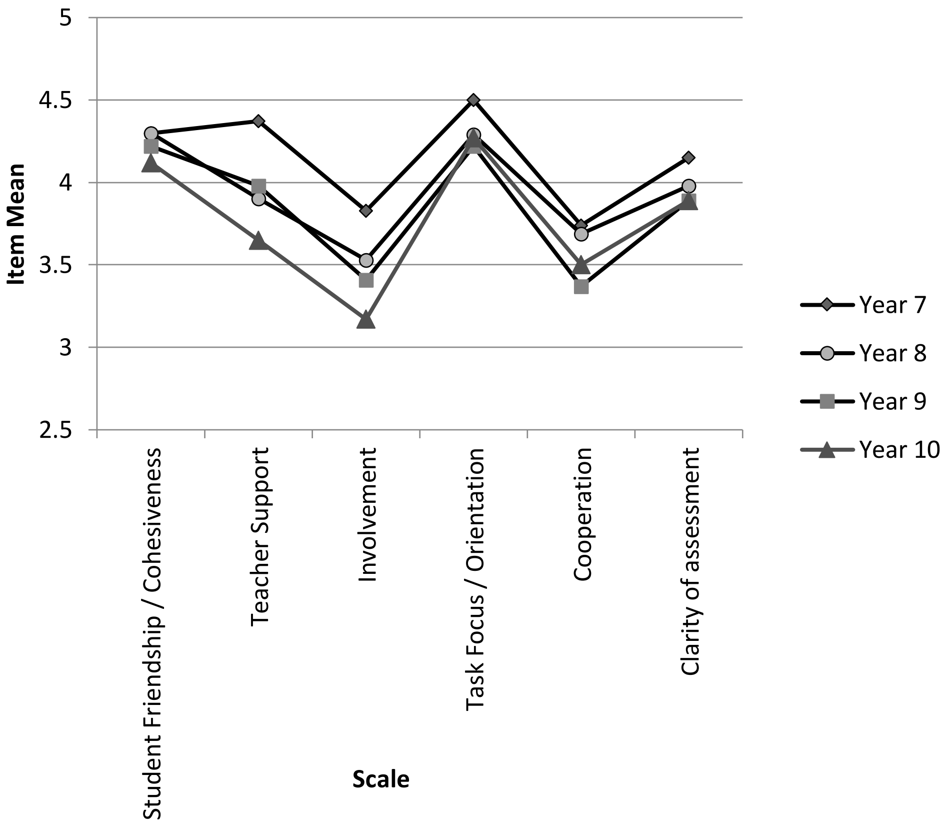

Figure 4.4 Item mean for actual WIHIC for students in Years 7, 8, 9 and 10

As in the previous section, the procedure used to compare group means was MANOVA. The six scales of the WIHIC (Student Friendship/Cohesiveness, Teacher Support, Involvement, Task Focus/Orientation, Cooperation and Clarity of Assessment Criteria) were used as the dependent variables with year level as the independent variable. The results of the MANOVA produced a Wilks' Lambda value of 0.803 with a probability value, $p < .05$, from which we can conclude that there is a statistically significant effect for year level. This suggests that there is a change in perception of the learning environment for different year levels. The final column of Table 4.12 gives the results of the MANOVA: the F value. The F value can be used to indicate whether there are statistically significant differences between the class means for each scale of the WIHIC. In this case it shows that there are statistically significant differences between year levels for three of the scales: Teacher Support ($p < .01$), Involvement ($p < .01$) and Cooperation ($p < .01$).

Tukey's HSD multiple comparisons procedure was carried out to ascertain whether there were statistically significant differences between particular pairs of year levels. The results are summarised in Table 4.13, in which asterisks indicate for which pairs of year levels there were statistically significant differences. The results indicate that there were statistically significant differences between some year levels for three scales. For Teacher Support (F= 9.65 and $p < .01$) there were statistically significant differences between: Years 7 and 8 ($p < .01$, effect size = 0.68); Years 7 and 9 ($p < .01$, effect size = 0.65); Years 7 and 10 ($p < .01$, effect size = 1.07); and Years 9 and 10 ($p < .05$, effect size = 0.44). For Involvement there were statistically significant (F = 9.65 and $p < .01$) differences between: Years 7 and 9 ($p < .05$, effect size = 0.55); Years 7 and 10 ($p < .01$, effect size = 0.81); and Years 8 and 10 ($p < .05$, effect size = 0.41). For Cooperation (F = 4.70 and $p < .01$) there were statistically significant differences between Years 7 and 9 ($p <.05$, effect size = 0.56) and Years 8 and 9 ($p < .05$, effect size = 0.48).

Table 4.12 Average item mean, average item standard deviation and MANOVA results for differences between Year 7, 8, 9 & 10 students' perceptions of the classroom learning environment measured using the WIHIC

Scale	Average Item Mean				Average Item Standard Deviation				F
	Year 7	Year 8	Year 9	Year 10	Year 7	Year 8	Year 9	Year 10	
Student Friendship/ Cohesiveness	4.30	4.30	4.22	4.12	0.61	0.59	0.48	0.70	1.50
Teacher Support	4.37	3.90	3.98	3.65	0.49	0.84	0.69	0.82	9.65**
Involvement	3.83	3.53	3.41	3.17	0.69	0.85	0.84	0.92	6.84**
Task Focus/ Orientation	4.50	4.29	4.22	4.27	0.42	0.58	0.64	0.67	2.27
Cooperation	3.74	3.69	3.37	3.50	0.66	0.68	0.66	0.78	4.70**
Clarity of Assessment Criteria	4.15	3.98	3.89	3.89	0.58	0.73	0.75	0.69	1.73

*$p < .05$ **$p < .01$

N = 45 students in Year 7, 81 students in Year 8, 102 students in Year 9, 97 students in Year 10.

Table 4.13 Effect size and Tukey's HSD multiple comparison for statistical significance of difference between each pair of year levels for the WIHIC

Scale	Effect Size[a] and Significance Level					
	Years 7 – 8	years 7 – 9	Years 7 – 10	Years 8 – 9	Years 8 – 10	Years 9 – 10
Student Friendship/ Cohesiveness	0	0.15	0.27	0.15	0.28	0.17
Teacher Support	0.68**	0.65*	1.07**	-0.10	0.30	0.44*
Involvement	0.39	0.55*	0.81**	0.14	0.41*	0.27
Task Focus/ Orientation	0.41	0.52	0.41	0.11	0.03	-0.08
Cooperation	0.07	0.56*	0.33	0.48*	0.26	-0.18
Clarity of assessment	0.26	0.39	0.41	0.12	0.13	0

*p < .05, ** *p* < .01
N = 45 students in Year 7; 81 students in Year 8; 102 students in Year 9; 97 students in Year 10.
[a] *the effect size is the difference between two means divided by the pooled standard deviation, providing the magnitude of the difference in terms of standard deviation units.*

The magnitudes of the differences between each pair of year levels for the three scales of the modified WIHIC, which show statistically significant differences between year levels, range from small to large as per the guidelines for eta^2 specified by J. Cohen (2013, pp. 25–26). Figures 4.5 to 4.7 show the effect sizes graphically. It can be seen from Figure 4.5 that the effect size between year levels for the Teacher Support scale are above medium for the differences between Years 7 and 8 and Years 7 and 9, large for Years 7 and 10, and between small and medium for Years 9 and 10.

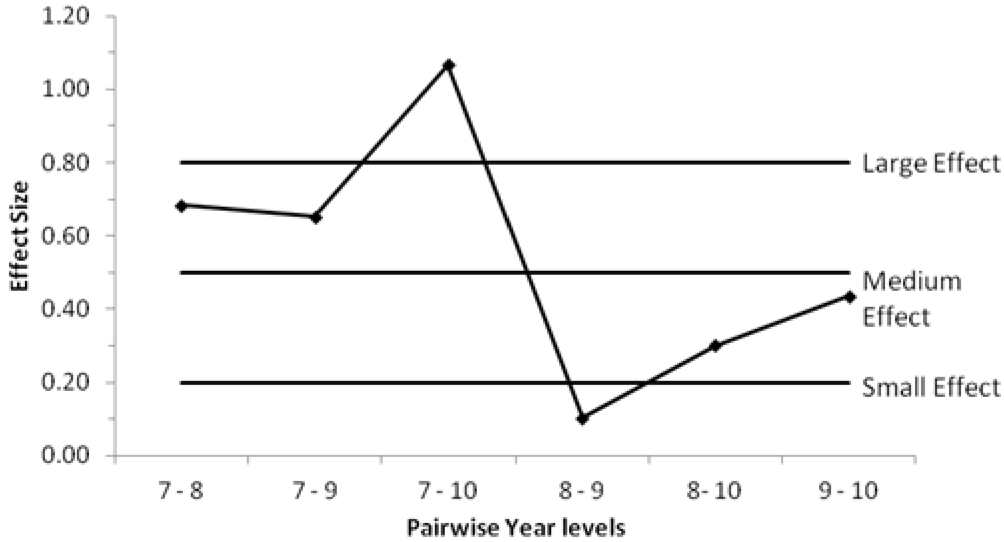

Figure 4.5 Effect size between pairs of year levels for the Teacher Support scale of the WIHIC

Figure 4.6 shows that for the Involvement scale the effect sizes range from between small and medium to large, with the differences between Years 7 and 9 having a medium effect size, Years 7 and 10 a large effect size and Years 8 and 10 a small to medium effect size.

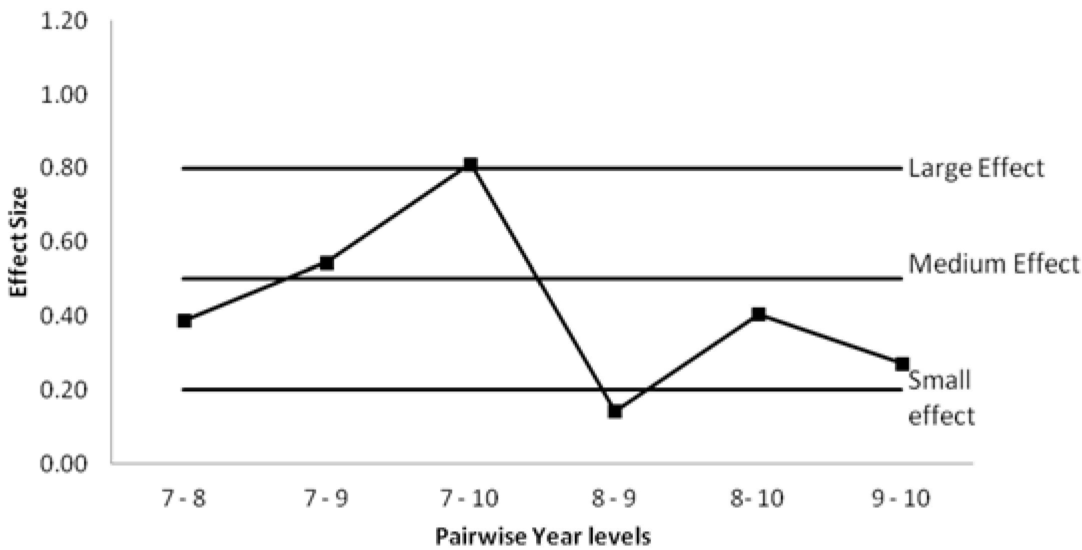

Figure 4.6 Effect size between pairs of year levels for the Involvement scale of the WIHIC

Figure 4.7 shows that for the Cooperation scale the effect sizes between Years 7 and 9 and Years 8 and 9 are just above and below medium, respectively.

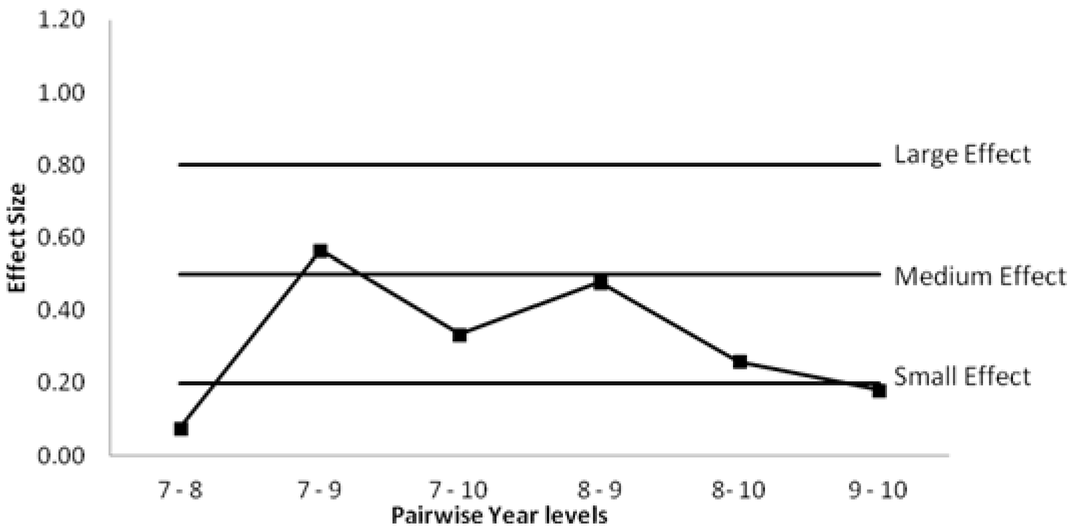

Figure 4.7 Effect size between pairs of year levels for the Cooperation scale of the WIHIC

4.5 Differences in Students' Confidence in Their Ability to Do and Learn Mathematics and Their Perceptions of the Classroom Learning Environment Based on Ability Group

The fourth objective of this study was to explore whether there are differences in students' confidence in their ability to do and learn mathematics and their perceptions of the learning environment based on whether they are in extension or mainstream groups. The data collected using the m-ATMI and the WIHIC were analysed to address the Research Question 4:

> Do students in extension classes differ from other students in the same year level in terms of their
>
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their learning environment?

The following sections report the results of the analyses undertaken to answer these questions. Students were in one of two groups: those who were in an extension class for some or all of their lessons ($N = 67$) and those who were not ($N = 268$). Section 4.5.1 records the results of the multivariate analysis of variance (MANOVA) used to determine whether there were differences between students' levels of confidence in their ability to do and learn mathematics based on whether they were in an extension class or not. The results of the MANOVA used to determine whether there are differences between students' perceptions of the learning environment, as measured by the WIHIC, based on group membership are recorded in Section 4.5.2.

4.5.1 Differences Between Student Confidence in Extension and Mainstream Classes

This section describes the results of the analysis carried out to determine whether there were differences in students' confidence in their ability to do and learn mathematics based on whether they were in an extension class or not. Figure 4.8 provides a graphical profile of the average item means (reported in Table 4.14) for each of the two groups (extension and mainstream). Whilst this profile does not provide information about whether these differences are statistically significant, inspection of the graphs indicates that there is a difference in means between each group for all of the scales of the m-ATMI. The standard deviations (reported in Table 4.14) are slightly higher for the mainstream group.

The procedure used to compare group means and to determine whether the differences shown in the graph are statistically significant was MANOVA. The four scales of the ATMI (Confidence, Value, Enjoyment and Motivation) were used as the dependent variables, with group membership (GroupID) as the independent variable. Group 1 is the extension group and Group 2 the mainstream group. The results of the MANOVA reported in the last column of Table 4.14 show that there were statistically significant differences between groups for all four scales: Confidence ($p < .01$); Value ($p < .01$); Enjoyment ($p < .01$); and Motivation ($p < .01$).

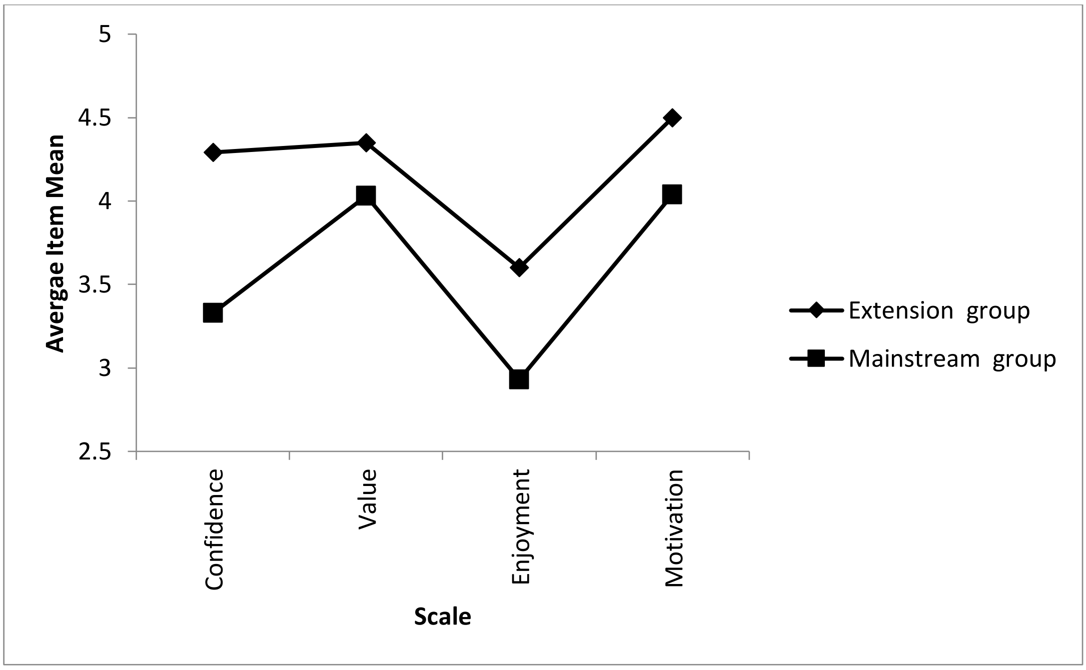

Figure 4.8 Means of the m-ATMI scales for extension and mainstream students

The effect size for each scale is recorded in Table 4.14. The asterisks in this table show the results of Tukey's HSD multiple comparisons procedure, which was carried out to ascertain whether differences were statistically significant. In this case the differences were statistically significant for all four scales of the m-ATMI. For Confidence (F = 68.01, p < .01), there was a statistically significant difference between the extension and mainstream groups (p < .01, effect size = 1.24). For Value (F = 8.48, p < .01), there was a statistically significant difference between the extension and mainstream groups (p < .01, effect size = 0.43). For Enjoyment (F = 29.34, p < .01), there was a statistically significant difference between the extension and mainstream groups (p < .01, effect size = 0.75). For Motivation (F = 23.00, p < .01), there was a statistically significant difference between the extension and mainstream groups (p < .01, effect size = 0.67).

The graph of effect sizes in Figure 4.9 shows that these range from between small to medium for the Value scale, between medium and large for the Enjoyment and Motivation scales and up to very large for the Confidence scale. These results suggest that students in extension classes had a better attitude towards mathematics, and in particular were more confident than students in mainstream classes.

Table 4.14 Average item mean, average item standard deviation and MANOVA results of differences between extension and mainstream groups using the m-ATMI

Scale	Average Item Mean		Average Item Standard Deviation		F	Effect Size[a] & Statistical Significance
	Extension classes	Mainstream group	Extension classes	Mainstream group		
Confidence	4.29	3.33	0.63	0.90	68.01**	1.24**
Value	4.35	4.03	0.71	0.82	8.48**	0.43**
Enjoyment	3.60	2.93	0.88	0.91	29.34**	0.75**
Motivation	4.50	4.04	0.66	0.71	23.00**	0.67**

*p < .05 **p < .01
N = 67 students in extension groups 268 students in mainstream groups

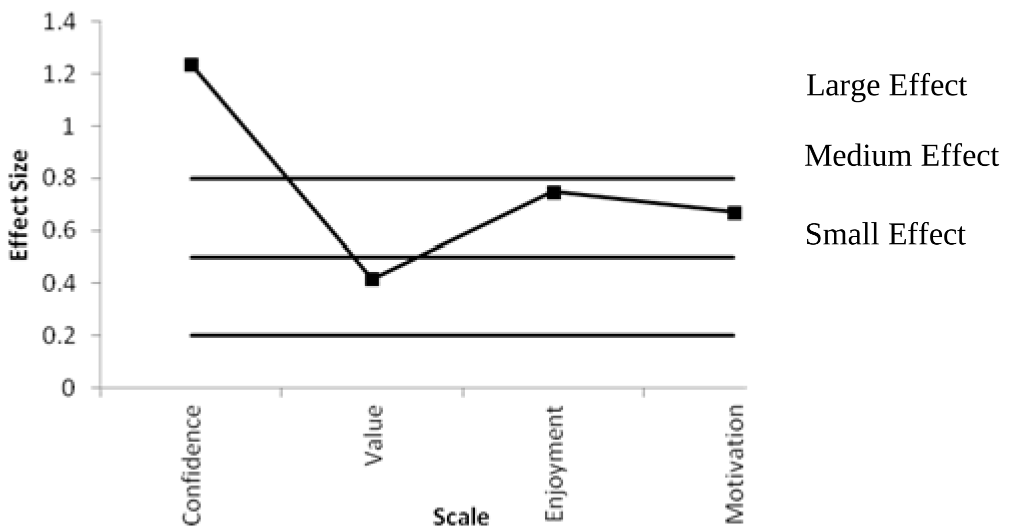

Figure 4.9 Effect Size between extension and mainstream students for m-ATMI

4.5.2 Differences Between Students' Perceptions of the Learning Environment in Extension and Mainstream Classes

Table 4.15 records the average item means and standard deviations for each group for each scale of the WIHIC. The graph in Figure 4.10 provides a visual profile of the average means for each scale of the WIHIC for each of the two groups of students (extension and mainstream). Inspection of the graph suggests that there are differences in the means for four of the seven scales (Involvement, Task Focus/Orientation, Cooperation and Clarity of Assessment Criteria), with students in the extension classes having higher average means for these items. The standard deviations for the mainstream group are slightly higher than those of the extension group. It is not possible to determine from the graph whether these differences are statistically significant.

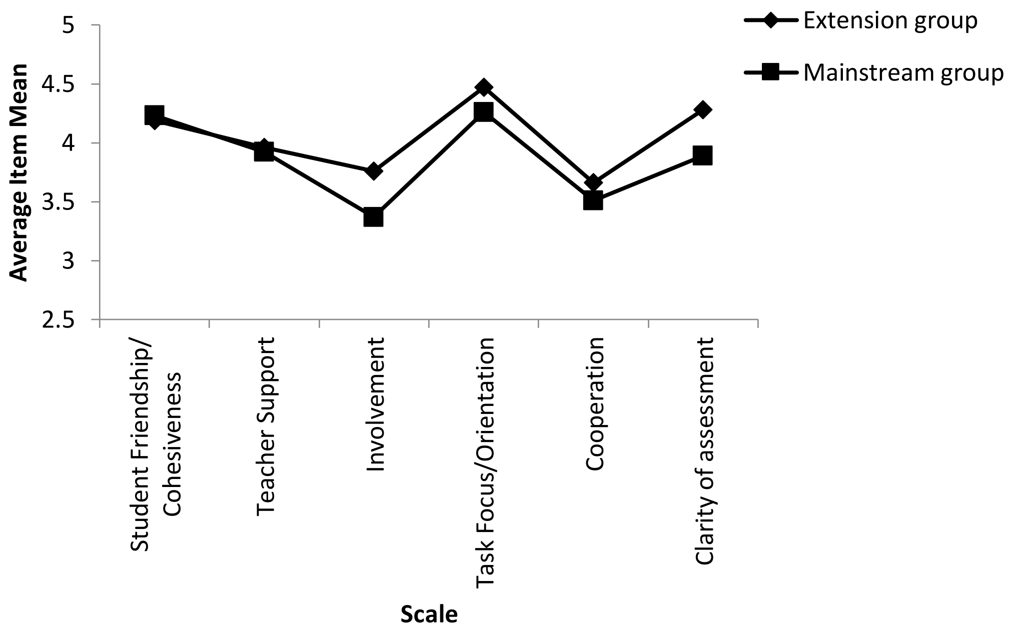

Figure 4.10 Graph of average item means of WIHIC scales for extension and mainstream groups

The procedure used to compare group means and hence determine whether the differences between the means are statistically significant was MANOVA. The six scales of the WIHIC (Student Friendship/Cohesiveness, Teacher Support, Involvement, Task Focus/Orientation, Cooperation and Clarity of Assessment Criteria) were used as the dependent variables, with group membership as the independent variable. The analysis produced a Wilks' Lambda score of 0.910 with $p < .05$, which suggests that there is a statistically significant difference between students' perceptions of the classroom learning environment based on group membership. The results of the MANOVA are reported in the last column of Table 4.15. They show statistically significant differences between the two groups for three of the scales: Involvement ($p < .01$) Task Focus/Orientation ($p < .01$) and Clarity of Assessment Criteria ($p < .01$).

Table 4.15 Average item mean, average item standard deviation and MANOVA results for differences between extension and mainstream groups for the scales of the WIHIC

Scale	Average Item Mean		Average Item Standard Deviation		F	Effect Size[a] & Statistical Significance
	Extension group	Mainstream group	Extension group	Mainstream group		
Student Friendship/ Cohesiveness	4.19	4.23	0.53	0.61	0.27	0.07
Teacher Support	3.96	3.92	0.75	0.78	0.15	0.05
Involvement	3.76	3.37	0.79	0.88	11.16**	0.47**
Task Focus/ Orientation	4.47	4.26	0.49	0.62	6.92**	0.38**
Cooperation	3.66	3.51	0.65	0.72	2.60	0.22
Clarity of Assessment Criteria	4.28	3.89	0.46	0.74	17.30**	0.63**

*$p < .05$ **$p < .01$

N = 67 students in extension groups 268 students in mainstream groups

[a] *the effect size is the difference between two means divided by the pooled standard deviation, providing the magnitude of the difference in terms of standard deviation units.*

The data in Table 4.15 give the effect size for the differences between each group for each scale and the results of Tukey's HSD multiple comparison, which shows that the differences between the extension and mainstream groups are statistically significant for three scales of the WIHIC: Involvement ($p < .01$, effect size 0.47); Task Focus/Orientation ($p < .01$, effect size 0.30); and Clarity of Assessment Criteria

($p < .01$, effect size = 0.63). Figure 4.11 provides a graphical view of the effect sizes for the differences between the two groups (students in extension and mainstream classes) for each scale of the WIHIC. This figure shows that for Involvement and Task Focus/Orientation scales the effect size is small to medium, and for the Clarity of Assessment Criteria scale it is medium to large as per Cohen's guidelines given in Section 4.4.1.

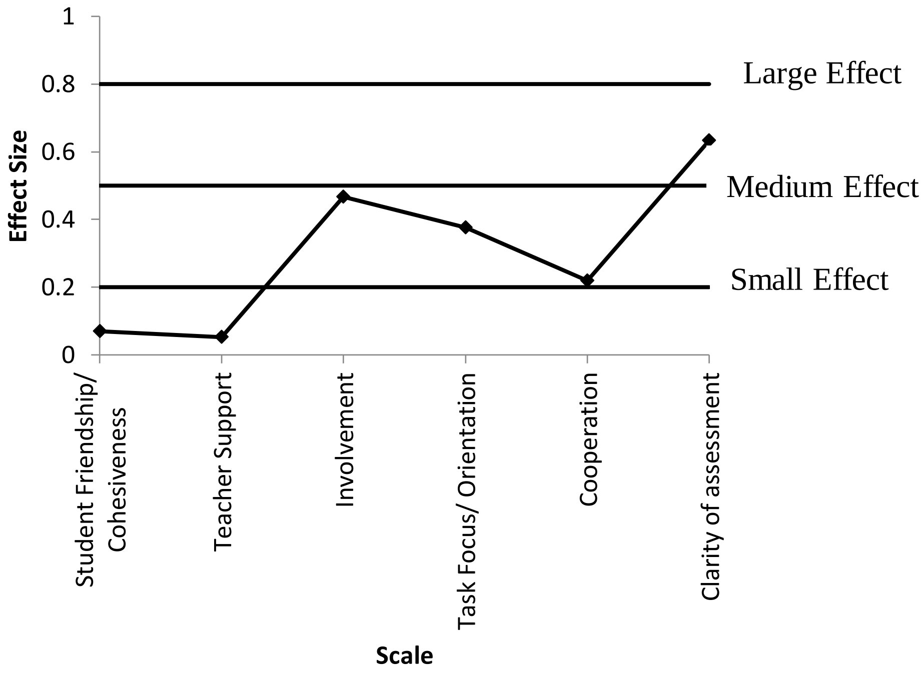

Figure 4.11 Effect size between extension and mainstream students for WIHIC scales

4.6 Differences Between Student Confidence in Different Strands of Mathematics

In School 1, each student completed both the m-ATMI and the WIHIC on three occasions. The timing of each administration of the questionnaires coincided with the completion of a unit of work, which it was hoped would come from different strands of mathematics. Due to absenteeism and the fact that some classes did not complete units of work from three different strands during the time of the study, the samples for this section of analysis were much smaller. The sample included thirty-five Year 8 students and two groups of Year 10 students (extension group N = 16 and mainstream group N = 15) who completed units of work from three different strands. MANOVA was used to analyse the data, the four scales of the m-ATMI were used as the dependent variables and the strand of mathematics was used as the independent variable. Because each group of students, as described above, completed units of work from different strands the analyses were run separately for each group. The results of the analyses are reported in Tables 4.16 to 4.19.

The results for the Year 8 students, who completed topics from three strands (Number, Algebra and Measurement), are reported in Table 4.16. The means and standard deviation for m-ATMI for each mathematics strand indicate that the differences between Confidence, Value and Motivation scales were quite small. In fact, the MANOVA indicated that there were no statistically significant differences between students' confidence levels for the three of the four m-ATMI scales. There was, however, a statistically significant difference between the means for the Enjoyment scale. Tukey's HSD multiple comparisons procedure was then carried out to ascertain whether there were statistically significant differences between particular pairs of mathematics strands. The results are summarised in Table 4.17, in which asterisks indicate pairs of strands for which there were statistically significant differences. The results indicated that for the Enjoyment (F = 25.31, and $p < .01$) scale there were statistically significant differences between the Algebra and Number strands ($p < .05$, effect size = 1.65) and the Algebra and Measurement strands ($p < .05$. effect size = 2.17). Inspection of the means suggests that in both cases, students reported higher levels of enjoyment for the Algebra (mean = 4.43) strand than for either the Number (mean = 3.01) or Measurement (3.14) strand. The effect sizes for these two differences were 1.65 and 2.17, respectively, both of which, according to J. Cohen (2013, p. 26) are large.

Table 4.16 Average item mean, average item standard deviation and MANOVA results for the difference between strands for Year 8 students using the m-ATMI

Scale	Average Item Mean			Average Item Standard Deviation			F
	Number	Algebra	Measurement	Number	Algebra	Measurement	
Confidence	3.55	3.58	3.59	1.18	1.14	1.10	0.01
Value	4.33	4.27	4.35	0.82	0.86	0.84	0.09
Enjoyment	3.01	4.43	3.14	1.07	0.58	1.04	25.32**
Motivation	4.42	4.34	4.39	0.58	0.54	0.72	0.05

*$p < .05$ **$p < .01$
N = 34 students in Number strand, 35 students in Algebra and Measurement strands

Table 4.17 Effect size and Tukey's HSD multiple comparison for statistical significance of difference between each pair of mathematics strands for the m-ATMI for Year 8 students

Scale	Effect Size[a] and Significance Level		
	Number – Algebra	Number – Measurement	Algebra – Measurement
Confidence	0.03	0.05	0.01
Value	0.07	0.03	0.13
Enjoyment	1.65*	0.17	2.17*
Motivation	0.14	0.06	0.11

*$p < .05$, ** $p < .01$
N = 34 students in Number strand and 35 students in the Algebra and Measurement strands
[a] the effect size is the difference between two means divided by the pooled standard deviation, providing the magnitude of the difference in terms of standard deviation units.

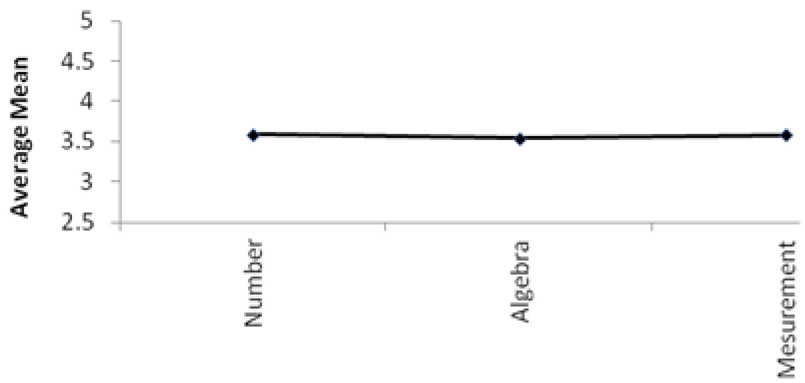

Figure 4.12 Average means of the Confidence scale for Year 8 students for three strands.

The Year 10 extension group completed the m-ATMI for three mathematics strands: Measurement, Algebra and Trigonometry. As with the Year 8 group, the mean scores for each of the m-ATMI scales, reported in Table 4.18, indicate that students' scores did not differ for most of the scales. The MANOVA results reported in the last column of Table 4.18 indicate that for the Year 10 extension group there were no statistically significant differences in students' confidence for the different strands of mathematics.

Table 4.18 Average item mean, average item standard deviation and MANOVA results for the difference between strands for Year 10 extension students using the m-ATMI

Scale	Average Item Mean			Average Item Standard Deviation			F
	Measurement	Algebra	Trigonometry	Measurement	Algebra	Trigonometry	
Confidence	4.03	3.86	3.92	0.50	0.50	0.60	0.41
Value	4.07	3.98	4.22	0.64	0.50	0.60	0.69
Enjoyment	3.28	3.19	3.26	0.72	0.68	0.87	0.07
Motivation	4.38	4.26	4.29	0.51	0.58	0.81	0.13

**p < .05 **p <.01*
N = 16 students in Measurement, Algebra and Trigonometry strands

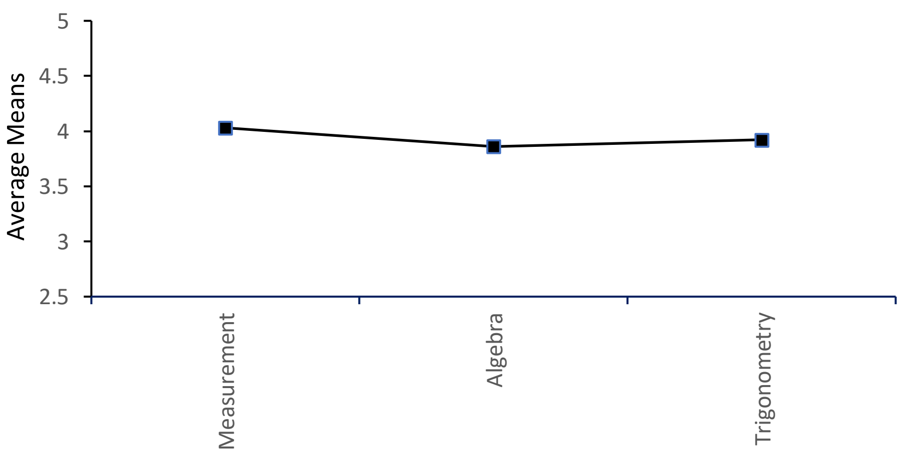

Figure 4.13 Average means of the Confidence scale for Year 10 extension students for three strands

The Year 10 mainstream group also completed three strands of mathematics, these being Measurement, Algebra and Statistics. The mean scores for each of the m-ATMI scales, reported in Table 4.19, suggest that the results for the three strands were relatively similar. The results of the MANOVA, recorded in the last column of Table 4.19, confirm this, as there were no statistically significant differences between the three strands for any of the m-ATMI scales.

Table 4.19 Average item mean, average item standard deviation and MANOVA results for the difference between strands for Year 10 mainstream students using the m-ATMI

Scale	Average Item Mean			Average Item Standard Deviation			F
	Measurement	Algebra	Statistics	Measurement	Algebra	Statistics	
Confidence	3.44	3.20	3.55	0.98	1.02	1.11	0.49
Value	3.78	3.82	3.85	0.85	0.74	0.81	0.03
Enjoyment	2.85	3.00	3.09	0.80	0.66	1.02	0.35
Motivation	4.48	4.28	4.21	0.43	0.52	0.64	1.15

**p < .05 **p <.01*
N = 17 students in Measurement strand, 16 students in Algebra and Statistics strands

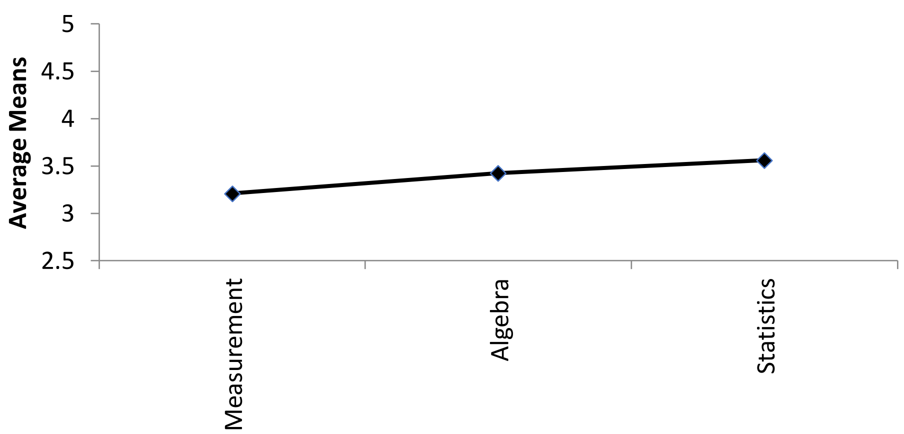

Figure 4.14 Average means of the Confidence scale for Year 10 mainstream students for three strands

4.7 Chapter Summary

This chapter has provided details of the analyses undertaken in order to answer the research questions posed in Chapter 1.

As a first step, analysis was undertaken to provide support for the reliability and validity of the instruments when used in South Australia. A number of techniques were employed to check the reliability and validity of the instruments used with 335 female students from 2 schools. Principal components analysis was used to check the factor structure of both the m-ATMI and the WIHIC. Item analysis of the m-ATMI maintained all scales and Cronbach's alpha provided evidence of good internal consistency reliability. The results reported in the component correlation matrix provided evidence of good discriminant validity. During item analysis of the WIHIC, the six items of the Equity scale and seven other items were found to be problematic and were removed from further analysis. The results provided strong support for the factor validity of the WIHIC with this sample. Cronbach's alpha confirmed that the WIHIC has a high internal consistency reliability for this sample. A one-way ANOVA performed for the actual version of the WIHIC provided evidence that four scales, namely Teacher Support, Involvement, Cooperation and Clarity of Assessment Criteria, were able to differentiate between students' perceptions in different classes. The results of the component correlation matrix suggest that there is good discriminant validity for the WIHIC.

To address the second research objective, relationships between confidence, classroom learning environment and achievement were explored using simple correlation and multiple regression techniques. The results of the simple correlation showed a statistically significant relationship between confidence and all the scales of the WIHIC with Involvement and Clarity of Assessment Criteria making the greatest contributions to the prediction of Confidence. It also showed that there is a statistically significant relationship between all scales of the m-ATMI and achievement. The results of the multiple correlation showed that 40% of the variance in Confidence is due to the learning environment model, while 32% of the variance in achievement is due to the m-ATMI model, with Confidence making the greatest contribution.

To address the third research objective, differences between students' confidence in their ability to do and learn mathematics and their perceptions of the learning environment based on their year level were examined using MANOVA. For students' responses to the m-ATMI, statistically significant differences were found for two scales – Value (between students in Years 7 and 10 and students in Years 8 and 10) and Enjoyment (between students in Years 7 and 10) – but not for Confidence. Statistically significant classroom learning environment differences were found for three of the six WIHIC scales (Teacher Support, Involvement and Cooperation). For the Teacher Support scale there were statistically significant differences between student responses for students in Years 7 and 8, Years 7 and 9, Years 7 and 10 and Years 9 and 10. For the Involvement scale, there were statistically significant differences between students' responses in Years 7 and 9, Years 7 and 10 and Years 8 and 10. Finally, for the Cooperation scale, there were statistically significant differences between students in Years 7 and 9 and Years 8 and 9.

To address the fourth research objective, MANOVA was also used to explore the differences in students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment based on their membership of either extension or mainstream classes. The results of the MANOVA showed a statistically significant difference between groups for all four m-ATMI scales (Confidence, Value, Enjoyment and Motivation) and for three of the six WIHIC scales (Involvement, Task Focus/Orientation and Clarity of Assessment Criteria). Students in the extension classes were more confident and had better perceptions of the classroom learning environment than those in mainstream classes.

Finally, to address the fifth research objective, MANOVA was calculated to determine whether there were statistically significant differences between students' confidence in their ability to do and learn mathematics for different strands of Mathematics. The results of the analyses, which were based on small samples of students in Years 8 and 10 (those who had completed topics tests in the same three strands) showed no statistically significant differences between these strands: Number, Algebra and Measurement for Year 8 students; Measurement, Algebra and Trigonometry for the Year 10 extension students; and Measurement, Algebra and Statistics for the Year 10 mainstream students. There were, however, statistically significant differences between the Number and Algebra strands and the Algebra and Measurement strands for Year 8 students on the Enjoyment scale. In both cases the results suggested that students enjoyed Algebra more than the other two strands.

The next chapter discusses the results and limitations of the present study, while making recommendations for future research. It also provides an outline of the significance of the study.

Chapter 5

SUMMARY, CONCLUSIONS AND RECOMMENDATION

5.1 Introduction

The study reported in this thesis used a correlational survey design to collect quantitative data. The data were used to address five research questions relating to female students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The sample consisted of 335 female students in 25 mathematics classes at two Catholic girls' schools in Adelaide, South Australia.

This chapter concludes the thesis by summarising and discussing the findings and their implications. This takes place in Section 5.2. Section 5.3 details the limitations of the study, and Section 5.4 offers some suggestions for further research in this area. Section 5.5 highlights the study's contributions to the field, while Section 5.6 provides some concluding remarks.

Theoretically, this study has made a distinctive contribution by bringing together two research fields: classroom learning environments and attitudes towards mathematics (in particular, confidence). A major contribution of this study has been to identify psychosocial elements of the classroom learning environment that are associated with girls' confidence in their ability to do and learn mathematics. In addressing the five research questions listed below, the study found that Involvement and Clarity of Assessment Criteria were positively and statistically significantly related to confidence levels. The study is also significant because it has focused on girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. It is one of only a few to have explored changes in students' perceptions of the classroom learning environment based on ability grouping, and to my knowledge is the only one to have considered differences in students' perceptions of their ability to do and learn mathematics based on the strand of mathematics being studied.

5.2 Summary and Discussion of the Findings

This section summarises and then discusses the major findings related to each of the research questions.

5.2.1 Research Question 1: Reliability and Validity of the Instruments Used

The first research question,

> Are the surveys to be used to collect data for the present study valid and reliable for use in girls' Catholic schools in metropolitan South Australia?

was concerned with whether the versions of the m-ATMI and the WIHIC questionnaires used in this study were reliable and valid when used with a sample of middle-school girls in Adelaide, South Australia. The data collected from 335 girls attending 2 Catholic girls' schools were analysed in various ways to provide evidence to support the reliability and validity of the instruments. The results are summarised and discussed for the m-ATMI in Section 5.3.1.1 and the WIHIC in Section 5.3.1.2.

5.2.1.1 Reliability and Validity of the m-ATMI

Principal axis factoring with oblique rotation was used to examine the internal structure of the 23-item, 4-scale m-ATMI. The internal consistency reliability was estimated using Cronbach's alpha coefficient. The component correlation matrix generated during oblique rotation was examined to determine the discriminant validity of the m-ATMI scales. The key findings are summarised below:

- All items of the m-ATMI met the criteria with loadings greater than .40 on their own scale and less than .40 on the other three scales. Therefore, all 23 items of the m-ATMI were retained.
- The total percentage of variance explained was 72%.
- The Cronbach's alpha reliability coefficient ranged from 0.89 (Motivation) to 0.94 (Enjoyment) with the individual as the unit of analysis, and from

0.85 (Motivation) to 0.96 (Value) with the class as the unit of analysis. These values confirmed that the questionnaire had high internal consistency reliability for this sample.

- Examination of the component correlation matrix showed that the correlation between the factors of the m-ATMI range from .42 to .62. These values were small enough, according to T. A. Brown's (2014, p. 116) recommendation, to have adequate discriminant validity. Although these figures indicated some overlap, the factor analyses supported the independence of the factor scores for the m-ATMI.

The findings above compare favourably with the results of the original ATMI (Tapia & Marsh, 2004) and with more recent studies in a number of age groups, languages, and countries such as Australia (Majeed et al., 2013), Singapore (Lim & Chapman, 2013); Ghana (Asante, 2012); Taiwan (Lin & Huang, 2014); and the UAE (Afari, 2013; Khine & Afari, 2014). Most of these studies supported the four-factor structure of the ATMI and all reported satisfactory Cronbach's alpha reliability values. i.e. $\alpha > .70$. Lim and Chapman (2013) and Afari (2013) found that both three- and four-factor models fitted their data, with the three-factor model providing a better fit. These results provide evidence to support the reliability and validity of the m-ATMI with middle-school girls in Adelaide.

5.2.1.2 Reliability and Validity of the WIHIC

Principal axis factoring with oblique rotation was also used to examine the internal structure of the 42-item modified WIHIC (conducted separately for the actual and preferred versions). The internal consistency reliability was estimated using Cronbach's alpha coefficient, separately for the actual and preferred versions. An ANOVA for each WIHIC scale, with class membership as the independent variable, was performed to determine the ability of the scales to differentiate between classes. The component correlation matrix generated during oblique rotation was examined to determine the discriminant validity of the WIHIC scales. The key findings are summarised below:

- During item analysis, the factor loadings for the Equity scale and seven other items were found not to meet the criteria (to have a loading of at least .40 on it's a *priori* scale and less than .40 on the other WIHIC scales) and were removed from all further analyses. Of the remaining items, all but one item met this criteria, the exception being Item 8 for the Teacher Support scale for the preferred version.
- The total percentage of variance explained was 59.94% for the actual version of the WIHIC and 59.43% for the preferred version.
- With the students as the unit of analysis, the highest Cronbach's alpha coefficient was .85 for the Teacher Support scale and the lowest .70 for the Student Friendship/Cohesiveness scale of the actual version. With class as the unit of analysis the scores ranged from .92 for the Involvement scale to .71 for the Student Friendship/Cohesiveness scale. Therefore the internal consistency reliability was satisfactory for the actual and preferred versions of the WIHIC.
- The results of the ANOVA showed statistically significant differences ($p < .01$) for four of the six scales: Teacher Support, Involvement, Cooperation and Clarity of Assessment Criteria. This suggested that these four scales were able to differentiate between students' perceptions in different classes.
- Examination of the component correlation matrix showed that values ranged from .10 to .41 for the actual version and from .20 to .35 for the preferred version for different WIHIC scales. These scores are well below the .80 recommended by T. A. Brown (2014, p. 116), indicating that the six retained scales of the WIHIC had adequate discriminant validity.

The results of the present study replicate much past research which has supported the reliability and validity of the WIHIC when used in other studies in Australia (Aldridge et al., 1999; Dorman et al., 2003; Velayutham et al., 2011). The findings provide evidence to support the WIHIC as a reliable and valid questionnaire when used with this sample of middle-school girls in South Australia, which gives confidence in the results obtained for the other questions in the study.

5.2.2 Research Question 2: Relationships Between Confidence, Classroom Learning Environment and Achievement

The second research question,

> Do relationships exist between students' confidence in their ability to do and learn mathematics and their
>
> a. perceptions of their classroom learning environment?
> b. achievement scores in mathematics?

was concerned with examining the relationships between girls' confidence in their ability to do and learn mathematics and two key factors: their perceptions of the classroom learning environment and their achievement. This section reports the results of the simple and multiple correlations performed in order to answer this question. First associations between confidence and classroom learning environment are discussed in Section 5.2.2.1, and then associations between confidence and achievement are discussed in Section 5.2.2.2.

5.2.2.1 Associations Between Confidence and Perceptions of the Classroom Learning Environment

This section summarises the findings of the data analyses used to investigate the relationships between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. Simple correlations were used to investigate the bivariate relationships between the scales of the m-ATMI and the WIHIC. The results of interest to this study are those showing a relationship between the confidence scale of the m-ATMI and each scale of the WIHIC. The key findings are summarised below:

- For the simple correlation analysis, confidence was positively and statistically significantly ($p < .01$) related to all six WIHIC scales.
- The correlations were all in the medium and large range (J. Cohen, 2013, p. 80). Task Focus/Orientation (.38) had a medium correlation with

Confidence, while Involvement (.51) and Clarity of Assessment Criteria (.56) had large correlations with Confidence.

- The multiple regression, *R*, between Confidence and the WIHIC scales was .63, positive and statistically significant ($p < .01$).
- R^2 was .40, indicating that 40% of the variance in confidence was due to the classroom learning environment as described by the six scales of the WIHIC.
- The beta values indicated that two of the six WIHIC scales, Involvement ($\beta = .36$) and Clarity of Assessment Criteria ($\beta = .43$), were significantly ($p < .01$) and independently associated with Confidence.

The results reported above support the view that girls' perceptions of the classroom learning environment have a positive impact on their confidence in their ability to do and learn mathematics. In other words, the more positive their perceptions of aspects of the classroom learning environment, the more confidence they will have in their ability to do and learn mathematics. These results support much other research which has found strong links between classroom learning environment and self-concept-like beliefs such as confidence (Kleitman & Gibson, 2011; Yang, 2015), mathematics self-esteem (Chionh & Fraser, 2009; Tran, 2012), self-concept (House, 1975; Trautwein et al., 2006) and academic self-efficacy (Aldridge & Fraser, 2008; Dorman, 2001; Dorman & Adams, 2004; Dorman et al., 2006; Dorman & Fraser, 2009). The positive relationships found in this study, coupled with the findings of past research, provide strong evidence of a link between improved student confidence and improved classroom learning environment.

The results summarised above found that the two scales of the WIHIC which had the highest statistically significant and positive correlations with confidence were Clarity of Assessment Criteria and Involvement. Clarity of Assessment Criteria measures the extent to which assessment criteria are made clear to the students; it provides students with information about "the criteria by which they are being assessed" (Bell & Aldridge, 2014). These results suggest that girls are more confident in their ability to do and learn mathematics when they understand the assessment criteria being used. This finding supports a number of studies that have reported the importance of girls understanding the role of assessment in the learning process as well as the actual assessment tasks: see, for example, O'Donovan, Price, and Rust (2004), who remind us that while assessment judgements should be consistent and reliable they should also be transparent and understood by students and teachers, and Mueller (2017, p. 272), who argues that for evaluation (assessment) to be effective “it must be regarded by teacher and learner alike as a continual and integral part of the learning process”. One study using five scales of the WIHIC and the Students’ Perceptions of Assessment Questionnaire (SPAQ) found the SPAQ scale Transparency (a similar scale to Clarity of Assessment Criteria) to be the most potent predictor of the self-concept-like construct Academic Efficacy (Dorman et al., 2006, p. 14).

With regard to Involvement, which refers to the extent to which students participate in the class, these results suggest that girls are more confident in their ability to do and learn mathematics when they perceive that they have opportunities to participate in the learning process. Hattie (2012, p. 112) stresses the importance of involvement in the learning process, saying that “learning requires the active involvement of the learner”. Other researchers studying learning environments have also found positive and statistically significant relationships between students perceptions of their involvement in the learning process and self-concept-like beliefs (Aldridge & Fraser, 2008; Dorman, 2001; Dorman & Fraser, 2009; Koul et al., 2011; Velayutham & Aldridge, 2013).

These findings suggest that teachers can indeed play a part in developing and nurturing student confidence, which many consider plays a critical role in girls’ participation in higher-level mathematics courses at school and beyond (Thomson et al., 2012; Watt, 2007; Watt et al., 2017). In other words, to improve girls’ confidence in their ability to do and learn mathematics, teachers can focus on creating a more positive classroom learning environment with particular emphasis on Clarity of Assessment Criteria and Involvement. In particular, to improve student confidence, teachers would do well to ensure that students have a clear understanding of the assessment criteria being used to measure their achievement, and include activities that encourage and enable them to actively participate in their own learning *(Recommendation 1).*

5.2.2.2 Associations Between Confidence and Achievement

Simple correlation and multiple regression analyses were also performed to examine the relationships between the confidence scales of the m-ATMI and student achievement. Achievement was measured using common topic tests set by the teachers in each school. The results of interest to this study are those showing relationships between confidence and achievement. The key findings are summarised below:

- For the simple correlation analysis, positive correlations for all scales of the m-ATMI and achievement were statistically significant ($p < .01$).
- The correlations ranged from .26 for achievement and Value to .56 for achievement and Confidence. The correlation between confidence and achievement can be considered large according to J. Cohen (2013).
- The multiple regression, *R*, between Confidence and achievement was .57 and statistically significant $p < .01$.
- R^2 indicated that 32% of the variance in achievement was due to students' attitudes as described by the m-ATMI, with Confidence ($\beta = .54$, $p < .01$) making the greatest contribution.

The results of this study support past research that has reported strong positive relationships between achievement and confidence. It is important to differentiate between those studies that have used standardised tests to measure achievement (Fennema & Sherman, 1977, 1978; J. Lee, 2009; Tartre & Fennema, 1995; Thomson et al., 2017) and those that have used school marks (Nurmi et al., 2003; Yee, 2010). The latter group are more relevant to this study, which used topic tests to determine students' achievement scores. Whilst interesting, and to some extent intuitive, the positive relationships between confidence and achievement do not provide teachers with information about how they can improve student confidence or achievement because there is no indication of causation in the results. This study and the ones reported above have not examined the possibility of a reciprocal relationship between confidence and achievement, which is a likely possibility (Arens et al., 2017; Ganley & Lubienski, 2016; Huang, 2011; Marsh & Martin, 2011; Niepel et al., 2014; Schöber et al., 2018; Valentine et al., 2004). This point is elaborated on in the Limitations section.

The findings reported in the this section have identified a statistically significant relationship between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment and between confidence and achievement as measured by topic tests. The relationship between confidence and the classroom learning environment is positive and statistically significantly for six scales of the WIHIC, and supports the findings of other studies, suggesting that teachers can increase girls' confidence in their ability to do and learn mathematics by improving the classroom learning environment, with a particular emphasis on Clarity of Assessment Criteria and Involvement. Although the data shows a positive, statistically significant relationship between confidence and achievement, as do other studies, it is less clear what actions teachers can take in response, as there is not enough evidence to determine the direction of the relationship.

5.2.3 Research Question 3: Differences in Student Perceptions based on Year Level

The third research question,

> Do students in different year levels differ in terms of their
>
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their classroom learning environment?

was concerned with examining differences in girls' perceptions of their confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment based on their membership of different year levels. MANOVA was used to determine whether these differences existed. The differences in confidence in their ability to do and learn mathematics based on year level are reported in Section 5.2.3.1, and the differences in perceptions of the classroom learning environment based on year level are reported in Section 5.2.3.2.

5.2.3.1 Differences in Confidence Based on Year Level

In this section, the results of the data analysis used to investigate differences in students' confidence in their ability to do and learn mathematics based on their year level, are summarised. MANOVA was used to compare girls in four groups based on year level (Years 7, 8, 9 and 10) using the scales of the m-ATMI as the dependent variables. The key findings are summarised below:

- The average item means decreased across year levels for each scale.
- The differences between Years 7, 8, 9 and 10 students were statistically significant for the Value scale ($p < .01$), but not for the Confidence, Enjoyment or Motivation scales.
- There were statistically significant ($p < .05$) differences between Year 7 and 10 and Year 8 and 10 students for the Value scale.
- There were statically significant differences ($p < .05$) between Year 8 and 10 students for the Enjoyment scale.
- There were no statistically significant differences between year levels for the Confidence scale.

The findings reported above indicated that, although the means for students' confidence in their ability to do and learn mathematics decreased as they progressed through school, the differences were not statistically significant. The results support those of other studies (Jacobs et al., 2002; Nurmi et al., 2003; Thomson et al., 2017) which have found that confidence decreases as students progress through school. However, unlike Jacobs et al. (2002) and Nurmi et al. (2003), who reported that that these differences were statistically significant, this study found no statistically significant difference in the level of student confidence based on particular year level. The findings for this study show that the greatest decline in confidence occurred between Year 7 and Year 8, which in South Australia, at the time of the study, marked the transition from primary school to high school. Other studies (Midgley et al., 1989; Speering & Rennie, 1996; Wigfield et al., 1991) have also reported decreases in students' self-perceptions as they transition from primary school to high school.

Teachers need to be aware of the changes in girls' perceptions of their ability to do and learn mathematics as they progress through school, but particularly as they transition from primary school to high school. They need to ensure that they are creating classroom learning environments that will enable girls to at least maintain their former levels of confidence. A recommendation for further study could be to explore the reasons for this decline in confidence by conducting a longitudinal mixed methods study *(Recommendation 2)*.

5.2.3.2 Differences in Perceptions of the Classroom Learning Environment Based on Year Level

In this section, the results of the data analyses used to investigate differences in girls' perceptions of the classroom learning environment, based on year level, are summarised and discussed. MANOVA was used to compare students in four groups based on year level (Years 7, 8, 9 and 10) for a number of dependent variables, namely the scales of the WIHIC. The key findings are summarised below:

- For all six scales of the WIHIC there was a decrease in the average item means as students moved to higher year levels.
- On a pairwise year level basis there were statistically significant differences for Teacher Support between students in Years 7 and 8, Years 7 and 9, Years 7 and 10 and Years 9 and 10.
- There were statistically significant differences for Cooperation between students in Years 7 and 9 and Years 8 and 9.
- There were statistically significant differences for Involvement between students in Years 7 and 9, Years 7 and 10 and Years 8 and 10.

The results summarised above found that in general, older students' perceptions of the classroom learning environment were less favourable than those of their younger counterparts. Studies of differences in perceptions of classroom learning environments due to year level have reported mixed results. Randhawa and Michayluk (1975) reported lower scores on the scales of the LEI for Year 11 students compared with Year 8 students, whereas Dorman and Fraser (2009) found that scores increased for 7 of the 10 scales of the TROFLEI for students in Years 8–10, 11 and 12. Yang (2015) found that Year 9 students generally viewed mathematics classes less favourably than Year 7 and 8 students. Koul et al. (2011) reported that students perceptions of the classroom learning environment differ according to year levels, and D. B. Byrne et al. (1986) reported differences in scores for different scales in different year levels. They surmised that rather than declining or improving, students' perception of the classroom learning environment changes as they mature, with older students preferring more independence and younger students seeking more structure and direction. One reason for the decline in students' perceptions of the classroom learning environment reported in this study may be that in South Australia, at the time of the study, Year 7 students were included in primary schools, where students have greater contact with one teacher. Primary school classrooms tend to be generalist in nature compared with secondary school classrooms, which are more subject-specific. These differences result in students spending the greater part of the day with the same teacher, which will affect their relationship with the teacher and their classmates. Transition studies provide further evidence of the changes that take place as students move from one section of a school to another (Deieso, 2016; Feldlaufer et al., 1988; P. Ferguson & Fraser, 1998, 1999; Hine, 2001; Midgley et al., 1989; Telli et al., 2006).

Three scales of the WIHIC were identified, in the results reported above, as being statistically significantly different across year levels: Teacher Support, Cooperation and Involvement. Teacher Support, which refers to the extent to which the teacher relates to the students, showed differences between Year 7 students and all other year levels as well as between Year 9 and 10 students. These results suggest that girls feel less supported by their teachers as they move through middle school and in particular as they transition to high school. The results support those reported by a number of studies (Barber & Olsen, 2004; Feldlaufer et al., 1988; P. Ferguson & Fraser, 1998, 1999; Hine, 2001; Koul et al., 2011; Midgley et al., 1989; Speering & Rennie, 1996; Way, Reddy, & Rhodes, 2007; Yang, 2015). This is of particular significance to the teachers of girls because girls have reported "that a good teacher/student relationship is directly related to positive class and subject attitudes" (P. Ferguson & Fraser, 1998, p. 395).

The second scale, Cooperation, which refers to the extent to which students cooperate rather than compete, showed differences between Year 7 and 9 and Year 8 and 9 students. These results suggest that Year 9 girls perceive the classroom learning environment as providing fewer opportunities to work together and learn cooperatively. This is despite the fact that Koul et al. (2011) reported that Year 8 and 9 students had significantly higher perceptions of cooperation in their classes than Year 7 students, while Year 10 students' perceptions of the level of cooperation were significantly less.

Involvement, which refers to the extent to which students participate in the class, showed statistically significant differences between students in Years 7 and 9, Years 7 and 10 and Years 8 and 10. These results suggest that girls in higher year levels perceive that they have fewer opportunities to participate in the learning process than in previous year levels. Other studies have reported similar results (Deieso, 2016; Hine, 2001; Koul et al., 2011; Yang, 2015).

Feldlaufer et al. (1988) reported that students have fewer opportunities for cooperation and involvement in the learning process following the transition from primary to secondary school. Teachers need to be in tune with the needs of their students and be aware of the many changes that they face as they progress through school, in particular as they transition from primary to secondary school. Their role is to ensure that they create classroom learning environments that take account of the changing maturity levels of their students. They should be providing more structure for younger students and more independence for older students. Due to the changing nature of students' perceptions of the classroom learning environment over time, action research is recommended to provide teachers with more information about the types of classroom learning environments they need to create for each year level *(Recommendation 3).*

The results reported in this study found that girls' confidence in their ability to do and learn mathematics declined as they progressed through school, although the decline was not statistically significant. The results also suggest that girls' perceptions of their classroom learning environment decline as they progress through school, although this may in fact be a change in their perceptions of different aspects of the classroom learning environment due to their increasing maturity. The change in confidence levels may also be related to the students' changing perceptions of the classroom learning environment. Action research studies which focus on how changing aspects of the classroom learning environment may provide further confirmation of its role in changing girls' confidence in their ability to do and learn mathematics *(Recommendation 4).*

5.2.4 Research Question 4: Differences in Student Perceptions Based on Ability Group

The fourth research question,

> Do students in extension classes differ from other students in the same year level in terms of their
>
> a. confidence in their ability to do and learn mathematics?
> b. perceptions of their classroom learning environment?

explored differences in students' perceptions based on membership of a particular ability group. MANOVA was used to determine whether there were differences in students' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment based on ability group, in this case membership of extension or mainstream groups. The differences in confidence based on ability group are reported in Section 5.3.4.1 and the differences in perceptions of the classroom learning environment based on ability group are reported in Section 5.3.4.2.

5.2.4.1 Differences in Confidence Based on Ability Group

This section summarises the results of the data analyses performed to determine differences in girls' confidence in their ability to do and learn mathematics based on the ability level of the mathematics class. MANOVA was used to compare students in two ability groups (extension vs mainstream) for a number of dependent variables, the scales of the m-ATMI. The key findings are summarised below:

- Means are higher across all m-ATMI scales for students in the extension groups.
- The differences are statistically significant ($p < .01$) for all four scales of the m-ATMI.
- The effect size for Confidence (1.24, $p < .01$) is very large.

It is therefore appropriate to say that for this sample the students in extension classes were more confident than those in mainstream classes. The topic of ability grouping in schools is contentious worldwide. Despite the apparent lack of evidence in its favour, it has been widely adopted. Most studies have reported that while there are often benefits both in terms of achievement and confidence for those in high ability groups, this is offset by negative outcomes for students in other groups (Belfi et al., 2011; Herrmann et al., 2016; Houtte et al., 2012; Ireson & Hallam, 2009). Belfi et al. (2011), Francis et al. (2017) and Ireson and Hallam (2009) point out that the students' positive perceptions could be a result of the students being placed in that particular group.

Teachers and schools need to be aware that there are differences in students' confidence in their ability to do and learn mathematics based on the ability level of their mathematics class. Francis et al. (2017) and Belfi et al. (2011) found that ability grouping was only beneficial for stronger students. They argued that the impact of grouping students by ability on self-confidence was one explanation for the poorer achievement of lower ability groups. Hattie (2002, p. 462) argued that there is little benefit to students from ability grouping and that teachers would be better to focus on "good teaching" and providing an appropriate classroom learning environment. Schools therefore need to decide whether dividing students in this way is appropriate. Forgasz (2010), Jaremus, Gore, Fray, and Prieto-Rodriguez (2020) and Zevenbergen (2003) point out that one effect of ability grouping is that it may limit students' choice of mathematics courses at Year 10 and beyond. This study was hoping to find ways of increasing girls' participation in higher-level mathematics courses at school, and subsequently in STEM careers. It seems possible that ability grouping may in fact work against this goal. According to Forgasz (2010) and Johnston and Wildy (2016), research into ability grouping in Australia has focussed on the students in high ability groups. Therefore, one recommendation is for further study to investigate the impact of ability grouping on the confidence and achievement levels of students, especially girls, in low and middle ability groups *(Recommendation 5).*

5.2.4.2 Differences in Perceptions of the Classroom Learning Environment Based on Ability Group

This section reports the results of the data analyses performed to determine whether there are differences in students' perceptions of the classroom learning environment based on ability grouping. MANOVA was used to compare students in two groups based on ability level (extension vs mainstream) for a number of dependent variables, the scales of the WIHIC. The key findings are summarised below:

- Average means were higher for students in extension classes for all WIHIC scales.
- The differences were statistically significant ($p < .01$) for the Involvement, Task Focus/Orientation and Clarity of Assessment Criteria scales.
- Clarity of Assessment Criteria made the largest contribution to the variation in the perception of the classroom learning environment.

The results reported in this study indicated that the girls in extension groups had a better perception of their classroom learning environment than the girls in mainstream classes. They support the results reported by other studies. Quek et al. (2005a, 2005b), Rita and Martin-Dunlop (2011) and Peer and Fraser (2015) all reported statistically significant differences between perceptions of the classroom for students from different ability groups, with more able students generally having a better perception of the classroom learning environment.

The results summarised above found that there were statistically significant differences between ability groups for three scales of the WIHIC: Involvement, Task Focus/Orientation and Clarity of Assessment Criteria. The few studies that have examined differences in students' perceptions of the classroom learning environment based on ability groups have provided different information. Rita and Martin-Dunlop (2011) reported that gifted high school biology students had a more positive perception of the classroom learning environment than their non-gifted counterparts. The differences were significant for six of seven scales of the WIHIC, including Involvement and Task Focus/Orientation. These results suggest that students in high ability classes are more likely to complete set work, pay attention in class, stay on task in lessons and set goals.

Peer (2011) and Peer and Fraser (2015) studied 1,081 primary school students in 55 classes in 4 schools in Singapore; they reported that students in the Gifted Education group perceived the classroom learning environment differently from students in the

High Ability group, with statistically significant differences for the Involvement, Cooperation and Personal Relevance sales.

Fraser and Lee (2009) and S. Lee et al. (2003) studied 439 South Korean students in 13 classes divided into three ability groups based on science ability/preference. They reported that the students in the high ability group perceived the classroom learning environment more favourably than those in the middle and low groups. Cooperation and ability to self-study were the major differences.

I was not able to find any studies that included Clarity of Assessment Criteria as a scale of the learning environment instrument. Given that all the studies cited above found that students in high ability groups had more positive attitudes towards their classroom learning environment it is not surprising that these students also have a better perception of how relevant assessment is and how clearly it is explained to them.

The results supported above show that the girls in the higher ability group had better perceptions of the classroom learning environment than those in the lower group. More research needs to be conducted to provide data on the impact of ability grouping on students' perceptions of the classroom learning environment across all ability groups *(Recommendation 6)*.

The findings reported in this section also suggest that the girls in the extension mathematics classes were more confident in their ability to do and learn mathematics and had a better perception of the classroom learning environment than those in mainstream mathematics classes. The message for teachers, if they are going to use ability grouping, is that they need to work on improving less able girls' confidence and perceptions of the classroom learning environment by ensuring that their classrooms are places where girls can be more involved and focussed on their work and that they have a clear understanding of the assessment criteria in use. More research should be conducted into the suitability of using ability groups in order to determine its impact on all students not just the more able *(Recommendation 7)*.

5.2.5 *Research Question 5: Differences in Students' Perceptions Based on Strand of Mathematics Studied*

The final research question,

> Do confidence levels differ for different strands of mathematics?

was interested in exploring whether girls display different levels of confidence in their ability to do and learn mathematics based on the strand of mathematics being studied. MANOVA was performed on the data for three different groups of students. The selection of the groups was determined by which students had completed the questionnaire three times and had completed topics from three different strands. The groups included a Year 8 class, a Year 10 extension class and a Year 10 mainstream class. These classes had completed topics from the Number, Algebra and Measurement strands, the Algebra, Measurement and Trigonometry strands and the Algebra, Measurement and Statistics strands, respectively. The key results are summarised below:

- For all three classes, there was very little difference between the means for the Confidence scale.
- The differences in the means were not statistically significant for the Confidence scale.
- For the Year 8 class there was a statistically significant difference between the Number and Algebra strands for the Enjoyment scale ($p < .01$).

Although the sample sizes were small, the results suggest that for these participants, there was no difference in their confidence in their ability to do and learn mathematics based on the strand of mathematics being studied. The differences shown on the Enjoyment scale between the Number and Algebra strands suggest that when girls say that they are more confident, they may in fact be saying that they enjoy the Algebra strand more than the Number strand. I was not able to find any similar research with which to compare the above results.

This section has reported on the findings of the data analyses carried out to answer the five research questions presented in Chapter 1. The analysis has found the two instruments used, the m-ATMI and the WIHIC, to be reliable and valid for use with the sample used in this study. Positive associations between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment as described by the WIHIC and between confidence and achievement have been reported, with Clarity of Assessment Criteria making the most significant contribution to the variance in confidence. Although girls' confidence in their ability to do and learn mathematics declined as they progressed through school, there were no statistically significant differences found. Significant differences in girls' perceptions of their confidence in their ability to do and learn mathematics were reported based on membership of extension or mainstream mathematics classes. Girls in extension mathematics classes were reported to be more confident than those in mainstream classes. No statistically significant difference in girls' confidence in their ability to do and learn mathematics was reported based on which strand of mathematics was being studied. Statistically significant differences were reported for girls' perceptions of the classroom learning environment based on membership of particular groups. Girls in lower year levels and those in extension mathematics classes had better perceptions of the classroom learning environment.

5.3 Limitations of the Study

The current study, like all research, involves a number of limitations which need to be considered when interpreting the results. Those limitations are discussed in this section.

One of the goals of quantitative research is to be able to generalise the results. The ability to do this is determined by the quality of the sample: in other words, how representative the sample is of the population being studied. This research focussed on girls' confidence in their ability to do and learn mathematics. The sample for this study involved students drawn from two Catholic girls' schools. Therefore, generalisations to other education systems should be made with caution. It is recommended that future studies include schools from the other education sectors in South Australia, as well as country and regional schools (*Recommendation 8*).

The students who participated in this study were all attending girls' schools. Therefore, generalisations to all girls schools should be made with caution. It would be useful to do a similar study that included girls from coeducational schools, as the presence of boys in the classroom learning environment may be a factor in determining girls' confidence in their ability to do and learn mathematics (*Recommendation 9).*

Another factor determining the suitability of a sample is its size. For Research Question 5 in this study, the sample was small due to the nature of the data collection required. Students included in this sample needed to have completed the surveys three times throughout the school year and to have studied topics from the same three strands of mathematics. This was difficult to achieve, despite the fact that students who were absent on the day the survey was administered were encouraged to complete the online survey in their own time. It would therefore be unwise to generalise the results of the analysis, which found that there were no differences in girls' confidence in their ability to do and learn mathematics based on the strand of mathematics being studied. It is recommended that future studies be done using a larger sample (*Recommendation 10).*

This study used a quantitative research method, in particular correlational design. Two questionnaires were used to collect the data. These used closed questions to elicit responses from the participants. Closed questions are useful because they are easy to administer and the results are easy to collate and analyse. They also make it easier to compare respondents' answers, and they can be helpful in clarifying what a question is asking. In order to get a better understanding of confidence and its relationship to the classroom learning environment, the use of open-ended questions involving the nature of confidence would have been helpful. These questions could have been used to ascertain both teachers' and students' perceptions of what is meant by the word confidence. The responses could also have been used to refine the items of the confidence scale *(Recommendation 11).*

Anecdotal evidence suggested that students might feel more or less confident in their ability to do and learn mathematics based on the strand of mathematics being studied. The results of this study found no evidence to support this view, but did find that there was a difference based on students' perceived enjoyment of the strand. The use of open-ended questions would also have been helpful in determining differences and similarities in how students perceived confidence and enjoyment (*Recommendation 12)*.

Assessment data was based on topic tests written by the teachers in each school. As "No assessment of student learning is value free" (AAMT, 2017), there are likely to be a number of inconsistencies in the tests. This is a limitation when the results of the tests are used in the same way. A recommendation is that in future studies all students take the same test *(Recommendation 13)*.

Past research on the relationship between confidence and achievement has produced mixed results (Hattie, 1992; Huang, 2011). There is a growing body of work (Arens et al., 2017; Ganley & Lubienski, 2016; Huang, 2011; Marsh & Martin, 2011; Niepel et al., 2014; Schöber et al., 2018; Valentine et al., 2004) that supports the view that the relationship is in fact reciprocal. The current study did not consider this as part of its design. A recommendation for future study is to consider the reciprocal nature of the relationship between confidence and achievement (*Recommendation 14)*.

5.4 Summary of Recommendations

This section provides a summary of the recommendations made above.

Recommendation 1	Given the relationship between confidence and perceptions of the classroom learning environment, it is recommended that teachers create a classroom in which girls are fully aware of assessment requirements and are actively engaged involved in their own learning.
Recommendation 2	To provide data on the reasons for the decline in confidence as girls progress through school, a longitudinal study is recommended.
Recommendation 3	Due to the changing nature of girls' perceptions of the classroom learning environment over time, a longitudinal study is recommended. This will provide teachers with more

information about the types of classroom learning environments they need to create as students progress through school.

Recommendation 4	To provide further confirmation of the impact of the classroom learning environment on girls' confidence in their ability to do and learn mathematics, action research is recommended.
Recommendation 5	To provide data on the impact of ability grouping on the confidence of all students, not just those in high ability groups, research involving students of all ability levels is recommended.
Recommendation 6	To provide data on the impact of ability grouping on students' perceptions of the classroom learning environment, a mixed-methods study is recommended.
Recommendation 7	In the light of the results of Recommendations 5 and 6 together with consideration of the social justice issues associated with the practice, a reassessment of the use of ability grouping is recommended.
Recommendation 8	To enhance the transferability of results, it is recommended that future samples include girls from all education sectors and regions in South Australia.
Recommendation 9	To provide data on whether the presence of boys in classrooms makes a difference to girls' perceptions of the classroom learning environment, it is recommended that girls from coeducational schools and classes be included in future research.

Recommendation 10	To provide more reliable answers to the question of whether girls have different levels of confidence in their ability to do and learn mathematics for different strands of mathematics, a larger sample is recommended.
Recommendation 11	To provide more information about how students and teachers perceive confidence in one's ability to do and learn mathematics in order to refine the confidence scales, the use of open-ended questions is recommended.
Recommendation 12	To provide more data about girls' perceptions in relation to different strands of mathematics, the examination of other factors such as enjoyment is recommended.
Recommendation 13	To provide consistency in assessment, the use of the same test to determine the achievement grades for all students is recommended.
Recommendation 14	To provide more data about the relationship between confidence and achievement, a consideration of the reciprocal nature of their relationship is recommended.

5.5 Contributions of the Study

This section outlines the significance of the study reported in this thesis. It is, as far as I am aware, one of only a few studies within the field of learning environment research to examine the relationships between girls' perceptions of the classroom learning environment and their confidence in their ability to do and learn mathematics. As a consequence the findings of this study have the potential to address one of the main reasons cited for women's low rates of participation in STEM careers: lack of confidence. It is also one of a handful of studies that have explored differences in girls' perceptions of the classroom learning environment based on ability groupings and, to the best of my knowledge, only one other study has mentioned differences in confidence levels based on the strand of mathematics being studied.

The findings of the study, described above, provide insights into how to increase girls' confidence in their ability to do and learn mathematics, which is significant because lack of confidence has often been cited as an important factor in the retention of girls in higher-level mathematics classes and subsequent STEM careers (Forgasz et al., 2004; Ganley & Lubienski, 2016; Lubienski et al., 2013; Watt, 2006; Watt et al., 2006; Zeldin & Pajares, 2000). The results have determined that two aspects of the classroom learning environment (Clarity of Assessment Criteria and Involvement) are statistically significantly related to girls' confidence in their ability to do and learn mathematics. This information provides teachers with information about which aspects of the classroom learning environment they should promote in order to improve girls' confidence in their ability to do and learn mathematics.

The study found that while there is no statistically significant change in girls' confidence in their ability to do learn and learn mathematics as they progress through middle school, there is a change in their perceptions of the classroom learning environment. This information could be used by teachers to create more positive mathematics learning environments. The results suggest that to do so, they should focus on Teacher Support, Cooperation and Involvement.

This is also one of only a handful of studies which has used and validated a version of the ATMI (m-ATMI) and the WIHIC in South Australia. The study reported in this thesis administered both instruments and provided strong support for their reliability and validity when used with girls in middle school mathematics classes in South Australia, thus making them useful tools for teachers and researchers in assessing students' perceptions of the classroom learning environment and their attitudes towards mathematics. This means that they are suitable for use in evaluating initiatives designed to improve girls' confidence in their ability to do and learn mathematics, which will hopefully lead to their greater participation in higher-level mathematics and subsequent involvement in STEM careers.

The study is adding to the growing body of knowledge with regard to classroom learning environments and their associations with self-concept-like constructs, in this case confidence. It supports the findings of studies that have examined the associations between academic efficacy – another self-concept-like construct – and classroom learning environments, including Aldridge, Afari, et al. (2012), Dorman (2001), Dorman and Adams (2004), (Dorman et al., 2003), Dorman et al. (2006) and J. M. Ferguson and Dorman (2001).

5.6 Concluding Remarks

The research reported in this study has shown the m-ATMI and WIHIC to be reliable and valid for the sample involving students from two South Australian Catholic girls' schools, and that positive relationships exist between girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment. The results also showed a strong relationship between confidence and student achievement, although the exact nature of this relationship was not clear. Additionally, the study found differences in girls' confidence in their ability to do and learn mathematics and their perceptions of the classroom learning environment based on membership of particular groups.

The research in this study has contributed to the identification of factors within the classroom learning environment that could enhance girls' confidence in their ability to do and learn mathematics. This is important because lack of confidence is often given as a reason for girls' failure to pursue higher-level mathematics courses and STEM related careers. The findings have identified two main areas of the classroom learning environment through which teachers can influence girls' confidence in their ability to do and learn mathematics. The results suggest that a classroom that provides a positive learning environment, where the criteria and purpose of assessment are clear and where girls believe they have opportunities to participate in their own learning will produce girls with greater confidence in their ability to do and learn mathematics. An increase in girls' confidence will hopefully lead to their increased participation in higher-level mathematics courses at school and their subsequent involvement in STEM careers.

REFERENCE LIST

Afari, E. (2013). Examining the factorial validity of the attitudes towards mathematics inventory (ATMI) in the United Arab Emirates: Confirmatory factor analysis. *International Review of Contemporary Learning Research: An International Journal, 2*(1), 15–29. https://doi.org/10.12785/irclr/020102

Afari, E., Aldridge, J. M., Fraser, B. J., & Khine, M. S. (2013). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research, 16*(1), 131–150. https://doi.org/10.1007/s10984-012-9122-6

Akerman, T. (2016, Aug 24). Engineering a gender balance demands systemic support. *The Australian,* p. 32. Retrieved from https://search.proquest.com/docview/1813241251?accountid=10382

Aldridge, J. M., Afari, E., & Fraser, B. J. (2012). Influence of teacher support and personal relevance on academic self-efficacy and enjoyment of mathematics lessons: A structural equation modeling approach. *The Alberta Journal of Educational Research, 58*(4), 614–633.

Aldridge, J. M., & Fraser, B. J. (1997). *Examining science classroom environments in a cross-relational study.* Paper presented at the Proceedings Western Australian Institute for Educational Research Forum 1997.

Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research, 3*(2), 101–134. https://doi.org/10.1023/A:1026599727439

Aldridge, J. M., & Fraser, B. J. (2008). *Outcomes-focused learning environments: Determinants and effects*. Sense Publishers.

Aldridge, J. M., & Fraser, B. J. (2011). Monitoring an outcomes focused learning environment: A case study. *Curriculum Perspectives, 31*, 25–41.

Aldridge, J. M., Fraser, B. J., Bell, L. M., & Dorman, J. P. (2012). Using a new learning environment questionnaire for reflection in teacher action research. *Journal of Science Teacher Education, 23,* 259–290. https://doi.org/10.1007/s10972-012-9268-1

Aldridge, J. M., Fraser, B. J., & Huang, I. T.-C. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93*(1), 48–62. https://doi.org/10.1080/00220679909597628

Aldridge, J. M., Fraser, B. J., & Ntuli, S. (2009). Utilising learning environment assessments to improve teaching practices among in-service teachers undertaking a distance-education programme. *South African Journal of Education, 29,* 147–170. https://doi.org/10.15700/saje.v29n2a256

Anderson, G. J. (1970). Effects of classroom social climate on individual learning. *American Educational Research Journal, 7*(2), 135–152. https://doi.org/10.3102/00028312007002135

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & Vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621–634. https://doi.org/10.1037/edu0000163

Arnup, J. L., Murrihy, C., Roodenburg, J., & McLean, L. A. (2013). Cognitive style and gender differences in children's mathematics achievement. *Educational Studies, 39*(3), 355–368. https://doi.org/10.1080/03055698.2013.767184

Asante, K. O. (2010). Sex differences in mathematics performance among senior high students in Ghana. *Gender and Behaviour, 8*(2), 3279–3289. https://doi.org/10.4314/gab.v8i2.61947

Asante, K. O. (2012). Secondary students' attitudes towards mathematics. *Ife Psychologia*, *20*(1), 121–133.

Asghard, M. (1994). Science education in Brunei: A study of the relationship between classroom environment and student achievement. *Unicorn, 20*(3), 62–66.

Australian Association of Mathematics Teachers (AAMT). (2017). *AAMT position paper: The practice of assessing mathematics learning*. Maths Education. http://www.mathseducation.org.au/wp-content/uploads/2018/10/Assessment_position_paper_2017.pdf

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2010). *General Capabilities*. https://www.acara.edu.au/curriculum/general-capabilities

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research, 19*(1), 3–30. https://doi.org/10.1177/0743558403258113

Barrington, F., & Evans, M. (2016). *Year 12 mathematics participation in Australia - The last ten years*. Australian Mathematical Sciences Institute.https://amsi.org.au/publications/participation-in-year-12-mathematics-2006-2016/

Belfi, B., Goos, M., De Fraine, B., & Van Damme, J. (2011). The effect of class composition by gender and ability on secondary school students’ school wellbeing and academic self-concept: A literature review. *Educational Research Review, 7*(1), 62–74. https://doi.org/10.1016/j.edurev.2011.09.002

Bell, L. M., & Aldridge, J. M. (2014). Investigating the use of student perception data for teacher reflection and classroom improvement. *Learning Environments Research, 17*(3), 371–388. https://doi.org/10.1007/s10984-014-9164-z

Benabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics, 117*(3), 871–915. https://doi.org/10.1162/003355302760193913

Blickenstaff, J. C. (2005). Women and science careers: leaky pipeline or gender filter? *Gender and Education, 17*(4), 369–386. https://doi.org/10.1080/09540250500145072

Bong, M., & Skaalvik, E. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40. https://doi.org/10.1016/j.lindif.2009.05.004

Bray, J. H., & Maxwell, S. E. (1985). *Multivariate analysis of variance*. SAGE Publications Inc.

Brown, R. (1965). *Social Psychology*. The Free Press.

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research, Second Edition* (2nd ed.). Guilford Publications.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102*(4), 964–981. https://doi.org/10.1037/a0019644

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford University Press.

Burden, R. L., & Fraser, B. J. (1993). Use of classroom environment assessments in school psychology: A British perspective. *Psychology in the Schools, 30*(3), 232–240. https://doi.org/10.1002/1520-6807(199307)30:3%3C232::aid-pits2310300306%3E3.0.co;2-k

Burton, L. (1992, 17 August). *Implications of constructivism for achievement in mathematics* [Conference paper]. International Congress on Mathematical Education (ICME 7), Quebec.

Burton, L. (2004). Confidence is Everything? Perspectives of Teachers and Students on Learning Mathematics. *Journal of Mathematics Teacher Education, 7*(4), 357–381. https://doi.org/10.1007/s10857-004-3355-y

Byrne, B. M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept* (pp. 287–316). John Wiley & Sons, Inc.

Byrne, D. B., Hattie, J. A., & Fraser, B. J. (1986). Student perceptions of preferred classroom learning environment. *The Journal of Educational Research, 80*(1), 10–18. https://doi.org/10.1080/00220671.1986.10885714

Campanini, M. (2020). Giving girls the tools to succeed in STEM. *Independent Education, 50*(1), 16–17.

Catsambis, S. (2005). The gender gap in mathematics: Merely a step function? In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender Differences In Mathematics: An Integrative Psychological Approach* (pp. 220–245). Cambridge University Press.

Chionh, Y. H., & Fraser, B. J. (2009). Classroom environment, achievement, attitudes and self-esteem in geography and mathematics in Singapore. *International Research in Geographical and Environmental Education, 18*(1), 29–44. https://doi.org/10.1080/10382040802591530

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Taylor and Francis.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education* (8th ed.). Taylor and Francis. http://ebookcentral.proquest.com

Costello, R. W. (1986). Relationships among ability grouping, classroom climate and academic achievement in mathematics and English classes. In B. J. Fraser (Ed.), *The study of learning environments* (Vol. 3, pp. 60–67). Curtin University of Technology.

Cramer, R. J., Neal, T. M., & Brodsky, S. L. (2009). Self-efficacy and confidence: Theoretical distinctions and implications for trial consultation. *Consulting Psychology Journal: Practice and Research, 61*(4), 319–334. https://doi.org/10.1037/a0017310

Creswell, J. (2012). *Educational Research* (4th ed.). Pearson.

Cretchley, P. C. (2008). Advancing research into affective factors in mathematics learning: clarifying key factors, terminology and measurement. In M. Goos, R. Brown, & K. Makar (Eds.), *Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 147–153). MERGA.

Davadas, S. D., & Lay, Y. F. (2020). Contributing factors of secondary students' attitude towards mathematics. *European Journal of Educational Research, 9*(2), 489–498. https://doi.org/10.12973/eu-jer.9.2.489

Deieso, D. (2016). *Changes in learning environment and students' attitudes and anxiety associated with the transition from primary to secondary school mathematics.* [Unpublished doctoral dissertation]. Curtin University.

den Brok, P., Fisher, D., Rickards, T., & Bull, E. (2006). Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation, 12*(1), 3–25. https://doi.org/10.1080/13803610500392053

Di Martino, P., & Zan, R. (2010). ‘Me and maths’: Towards a definition of attitude grounded on students’ narratives. *Journal of Mathematics Teacher Education, 13*(1), 27–48. https://doi.org/10.1007/s10857-009-9134-z

Di Martino, P., & Zan, R. (2011). Attitude towards mathematics: A bridge between beliefs and emotions. *Mathematics Education, 43*, 471–482. https://doi.org/10.1007/s11858-011-0309-6

Di Martino, P., & Zan, R. (2015). The construct of attitude in mathematics education. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education: Exploring a mosaic of relationships and interactions* (pp. 51–72). Springer International Publishing.

Dillman, D. A. (2006). *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Retrieved from http://CURTIN.eblib.com.au/patron/FullRecord.aspx?p=281571

Dogan, H. (2012). Emotion, confidence, perception and expectation case of mathematics. *International Journal of Science and Mathematics Education, 10*(1), 49–69. https://doi.org/10.1007/s10763-011-9277-0

Dorman, J. P. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research, 4*(3), 243–257. https://doi.org/10.1023/a:1014490922622

Dorman, J. P. (2003). Cross-national validation of the What is Happening in this Class? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research, 6*(3), 231–245. https://doi.org/10.1023/A:1027355123577

Dorman, J. P. (2008). Use of multitrait-multimethod modelling to validate actual and preferred forms of the What is happening In this Class? (WIHIC) questionnaire. *Learning Environments Research: An International Journal, 11*, 179–193. https://doi.org/10.1007/s10984-008-9043-6

Dorman, J. P., & 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. https://doi.org/10.1080/0140672040270106

Dorman, J. P., Adams, J., & Ferguson, J. M. (2003). A cross-national investigation of students' perceptions of mathematics classroom environment and academic efficacy in secondary schools. *International Journal for Mathematics Teaching and Learning, 15*, 123–176.

Dorman, J. P., Fisher, D. L., & Waldrip, B. (2006). Classroom environment, students' perceptions of assessment, academic efficacy and attitude to science: A Lisrel analysis. In D. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: Worldviews* (pp. 1–28). World Scientific Publishing Co.

Dorman, J. P., & Fraser, B. J. (2009). Psychosocial environment and affective outcomes in technology-rich classrooms: Testing a causal model. *Social Psychology of Education: An International Journal, 12*(1), 77–99. https://doi.org/10.1007/s11218-008-9069-8

Druckman, D., & Bjork, R. A. (1994). *Learning, remembering, believing: Enhancing human performance*. National Academy Press.

Eccles, J. (2011a). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development, 35*(3), 195–201. https://doi.org/10.1177/0165025411398185

Eccles, J. (2011b). Understanding women's achievement choices: Looking back and looking forward. *Psychology of Women Quarterly, 35*, 510–516. https://doi.org/10.1177/0361684311414829

Eccles, J., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on Motivation in Education: Goals and Cognitions* (Vol. 3, pp. 139–186). Academic Press Inc.

Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS ONE, 11*(7), e0157447. https://doi.org/10.1371/journal.pone.0157447

Feldlaufer, H., Midgley, C., & Eccles, J. (1988). Student, teacher, and observer perceptions of the classroom environment before and after the transition to junior high school. *The Journal of Early Adolescence, 8*(2), 133–156. https://doi.org/10.1177/0272431688082003

Fennema, E. (1974). Sex differences in mathematics-learning: Why??? *The Elementary School Journal, 75*(3), 183–190. https://doi.org/10.1086/460891

Fennema, E. (2000, 22 May). *Gender and mathematics: What is known and what do I wish was known?* [Conference paper]. Fifth Annual Forum of the National Institute for Science Education., Detroit, Michigan, United States.

Fennema, E., & Carpenter, T. P. (1998). New perspectives on gender differences in mathematics: An introduction. *Educational Researcher, 27*(5), 4–5. https://doi.org/10.2307/1176732

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). New perspectives on gender differences in mathematics: A reprise. *Educational Researcher, 27*(5), 19–21. https://doi.org/10.2307/1176737

Fennema, E., & Sherman, J. (1976). Fennema-Sherman mathematics attitudes scales: instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324–326. https://doi.org/10.2307/748467

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal*, *14*(1), 51–71. https://doi.org/10.3102/00028312014001051

Fennema, E., & Sherman, J. (1978). Sex-related differences in mathematics achievement and related factors: a further study. *Journal for Research in Mathematics Education, 9*(3), 189–203. https://doi.org/10.2307/748997

Ferguson, J. M., & Dorman, J. P. (2001). Psychosocial classroom environment and academic efficacy in Canadian high school mathematics classes. *Alberta Journal of Educational Research, 47*(3), 276–279.

Ferguson, P., & Fraser, B. J. (1998). Student gender, school size and changing perceptions of science learning environments during the transition from primary to secondary school. *Research in Science Education, 28*(4), 387–397. https://doi.org/10.1007/bf02461506

Ferguson, P., & Fraser, B. J. (1999). Changes in learning environment during the transition from primary to secondary school. *Learning Environments Research, 1*(3), 369–383. https://doi.org/10.1023/A:1009976529957

Fisher, D. L., & Fraser, B. J. (1981). Validity and use of the my class inventory. *Science Education*, *65*(2), 145–156. https://doi.org/10.1002/sce.3730650206

Fisher, D. L., & Fraser, B. J. (1983). A comparison of actual and preferred classroom environments as perceived by science teachers and students. *Journal of Research in Science Teaching, 20*(1), 55–61. https://doi.org/10.1002/tea.3660200106

Fisher, D. L., Henderson, D., & Fraser, B. J. (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education, 25*(2), 125–133. https://doi.org/10.1007/bf02356447

Forgasz, H. (2010). Streaming for mathematics in Victorian secondary schools. *Australian Mathematics Teacher, 66*(1), 31–40.

Forgasz, H. (2012). *Towards equity in mathematics education: Gender, culture, and diversity*. Springer Berlin Heidelberg.

Forgasz, H., Leder, G., & Kloosterman, P. (2004). New perspectives on the gender stereotyping of mathematics. *Mathematical Thinking and Learning, 6*(4), 389–420. https://doi.org/10.1207/s15327833mtl0604_2

Francis, B., Connolly, P., Archer, L., Hodgen, J., Mazenod, A., Pepper, D., Sloan, S., Taylor, B., Tereshchenko, A., & Travers, M.-C. (2017). Attainment grouping as self-fulfilling prophesy? A mixed methods exploration of self confidence and set level among Year 7 students. *International Journal of Educational Research, 86*, 96–108. https://doi.org/10.1016/j.ijer.2017.09.001

Fraser, B. J. (1980). *Criterion validity of an individualised classroom environment questionnaire*. B. J. Fraser.

Fraser, B. J. (1991). Two decades of classroom environment research. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Education, antecedents and consequences* (pp. 3–28). Permagon Press.

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

Fraser, B. J. (1998a). The birth of a new journal: Editor's introduction. *Learning Environments Research, 1*, 1–5.

Fraser, B. J. (1998b). Classroom environment instruments: Development, validity and applications. *Learning Environments Research, 1*(1), 7–34. https://doi.org/10.1023/A:1009932514731

Fraser, B. J. (2001). Twenty thousand hours: Editor's introduction. *Learning Environments Research: An International Journal, 4*, 1–5.

Fraser, B. J. (2002). Learning environments research: Yesterday, today and tomorrow. In S. Goh & M. S. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 1–25). World Scientific Publishing

Fraser, B. J. (2012a). *Classroom Environment (RLE Edu O)*. Taylor and Francis.

Fraser, B. J. (2012b). Classroom learning environments: Retrospective, context and prospect. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *The second international handbook for science education* (pp. 1191–1239). 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. 2, pp. 229–253). Routledge.

Fraser, B. J. (2019). Milestones in the evolution of the learning environments field over the past three decades. In B. J. Fraser & D. B. Zandvliet (Eds.), *Thirty years of learning environments* (Vol. 2, pp. 1–19). Brill Sense.

Fraser, B. J., Aldridge, J. M., & Adolphe, F. S. G. (2010). A cross-national study of secondary science classroom environments in Australia and Indonesia. *Research in Science Education, 40*(4), 551–571. https://doi.org/10.1007/s11165-009-9133-1

Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)* (3rd ed.). Western Australian Institute of Technology.

Fraser, B. J., & Butts, W. L. (1982). Relationship between perceived levels of classroom individualization and science-related attitudes. *Journal of Research in Science Teaching, 19*(2), 143–154. https://doi.org/10.1002/tea.3660190206

Fraser, B. J., & Fisher, D. L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*(4), 498–518. https://doi.org/10.3102/00028312019004498

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1992). *Assessing the climate of science laboratory classes*. Curtin University of Technology, National Key Centre for School Science and Mathematics.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching, 32*(4), 399–422. https://doi.org/10.1002/tea.3660320408

Fraser, B. J., & Lee, S. (2009). Science laboratory classroom environments in Korean high schools. *An International Journal, 12*(1), 67–84. https://doi.org/10.1007/s10984-008-9048-1

Fraser, B. J., McRobbie, C. J., & Fisher, D. L. (1996, April). *Development, validation and use of personal and class forms of a new classroom environment questionnaire* [Conference paper]. Western Australian Institute for Educational Research Forum 1996, Bentley, Perth.

Fraser, B. J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary school classrooms. *The Elementary School Journal*, *85*(5), 567–580. https://doi.org/10.1086/461422

Fraser, B. J., & Tobin, K. (1991). Combining qualitative and quantitative methods in classroom environment research. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 271–292). Pergamon.

Fraser, B. J., & Treagust, D. F. (1986). Validity and use of an instrument for assessing classroom psychosocial environment in higher education. *Higher Education*, *15*(1), 37–57. https://doi.org/10.1007/bf00138091

Fraser, B. J., & Walberg, H. J. (1991). Introduction and overview. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents, and consequences* (pp. ix–xv). Pergamon Press

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics, 2*(2), 210–240. https://doi.org/10.1257/app.2.2.210

Galbraith, P., & Haines, C. (1998). Disentangling the nexus: Attitudes to mathematics and technology in a computer learning environment. *Educational Studies in Mathematics, 36*(3), 275–290. https://doi.org/10.1023/:1003198120666

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences, 47*, 182–193. https://doi.org/10.1016/j.lindif.2016.01.002

Goh, S. C., & Fraser, B. J. (1996). Validation of an elementary school version of the questionnaire on teacher interaction. *Psychological Reports, 79*(2), 515–522. https://doi.org/10.2466/pr0.1996.79.2.515

Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research, 1*(2), 199–229. https://doi.org/10.1016/j.tate.2015.04.007

Goodenow, C. (1993). Classroom belonging among early adolescent students. *The Journal of Early Adolescence, 13*(1), 21–43. https://doi.org/10.1177/0272431693013001002

Guskey, T. R. (2013). Defining student achievement. In J. Hattie & E. Anderman (Eds.), *International guide to student achievement* (pp. 3–6). Routledge.

Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education, 20*(3), 261–273. https://doi.org/10.2307/749515

Haertel, G. D., Walberg, H. J., & Haertel, E. H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7*(1), 27–36. https://doi.org/10.1080/0141192810070103

Harrington, M. (2011). *Australian government funding for schools explained*. Parliamentary Library.

Hart, L. E. (1989). Classroom processes, sex of student, and confidence in learning mathematics. *Journal for Research in Mathematics Education, 20*(3), 242–260. https://doi.org/10.2307/749514

Hattie, J. (1992). *Self-concept*. L. Earlbaum Associates.

Hattie, J. (2002). Classroom composition and peer effects. *International Journal of Educational Research, 37*(5), 449–481. https://doi.org/10.1016/s0883-0355(03)00015-6

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Taylor and Francis.

Helding, K. A., & Fraser, B. J. (2013). Effectiveness of National Board Certified (NBC) teachers in terms of classroom environment, attitudes and achievement among secondary science students. *Learning Environments Research, 16*(1), 1–21. https://doi.org/10.1007/s10984-012-9104-8

Herrmann, J., Schmidt, I., Kessels, U., & Preckel, F. (2016). Big fish in big ponds: Contrast and assimilation effects on math and verbal self-concepts of students in within-school gifted tracks. *British Journal of Educational Psychology, 86*(2), 222–240. https://doi.org/10.1111/bjep.12100

Hine, P. (2001). *Classroom environment and the transition to secondary schooling.* [Unpublished doctoral dissertation]. Curtin University.

House, P. A. (1975). Learning environments, academic self-concepts, and achievement in mathematics. *Journal for Research in Mathematics Education, 6*(4), 244–252. https://doi.org/10.5951/jresematheduc.6.4.0244

Houston, L. S., Fraser, B. J., & Ledbetter, C. E. (2008). An evaluation of elementary school science kits in terms of classroom environment and student attitudes. *Journal of Elementary Science Education, 20*(4), 29–47. https://doi.org/10.1007/bf03173675

Houtte, M. V., Demanet, J., & Stevens, P. A. (2012). Self-esteem of academic and vocational students: Does within-school tracking sharpen the difference? *Acta Sociologica, 55*(1), 73–89. https://doi.org/10.1177/0001699311431595

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505–528. https://doi.org/10.1016/j.jsp.2011.07.001

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*, 299–324. https://doi.org/10.1111/j.1471-6402.1990.tb00022.x

Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability group in schools. *Learning and Instruction, 19*(3), 201–213. https://doi.org/10.1016/j.learninstruc.2008.04.001

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

James, W. (1890). *The principles of psychology* (Vol. 53). Encyclopaedia Britannica.

Jansen, M., Schroeders, U., & Lüdtke, O. (2013). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences, 30,* 11–21. https://doi.org/10.1016/j.lindif.2013.12.003

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2019). The dimensional structure of students' self-concept and interest in science depends on course competition. *Learning and Instruction, 60,* 20–28. https://doi.org/10.1016/j.learninstruc.2018.11.001

Jaremus, F., Gore, J., Fray, L., & Prieto-Rodriguez, E. (2020). Grouped out of STEM degrees: the overlooked mathematics ‘glass ceiling’ in NSW secondary schools. *International Journal of Inclusive Education*, 1–17. https://doi.org/10.1080/13603116.2020.1776778

Jinks, J., & Lorsbach, A. (2003). Introduction: Motivation and self-efficacy belief. *Reading & Writing Quarterly, 19*(2), 113–118. https://doi.org/10.1080/10573560308218

Johnson, B., & McClure, R. (2004). Validity and reliability of a shortened, revised version of the constructivist learning environment survey (CLES). *Learning Environments Research, 7*(1), 65–80. https://doi.org/10.1023/b:leri.0000022279.89075.9f

Johnston, O., & Wildy, H. (2016). The effects of streaming in the secondary school on learning outcomes for Australian students: A review of the international literature. *The Australian Journal of Education, 60*(1), 42–59. https://doi.org/10.1177/0004944115626522

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education, 51*(4), 1609–1620. https://doi.org/10.1016/j.compedu.2008.03.003

Kennedy, J., Lyons, T., & Quinn, F. (2014). The continuing decline of science and mathematics enrolments in Australian high schools. *Teaching Science, 60*(2), 34–46.

Khan, A. K., & Rodrigues, S. (2012). The influence of school type on students' attitude towards mathematics in Quetta, Pakistan. *Journal of Research and Reflections in Education, 6*(2), 95–102.

Khine, M. S., & Afari, E. (2014). Psychometric properties of an inventory to determine the factors that affect students' attitudes toward mathematics. *Psychology, Society, & Education, 6*(1), 1–15. https://doi.org/10.25115/psye.v6i1.504

Khoo, H. S., & Fraser, B. J. (2008). Using classroom psychosocial environment in the evaluation of adult computer application courses in Singapore. *Technology, Pedagogy and Education, 17*(1), 67–81. https://doi.org/10.1080/14759390701847518

Khun-inkeeree, H., & Omar-fauzee, M. s. (2020). Original paper authenticating factor analysis of attitude towards mathematics inventory (ATMI) in Thailand. *Frontiers in Education Technology, 3*(1), 26–39. https:// doi.org/10.22158/fet.v3n1p26

Kim, H.-B., Fisher, D. L., & Fraser, B. J. (1999). Assessment and investigation of constructivist science learning environments in Korea. *Research in Science & Technological Education, 17*(2), 239–249. https://doi.org/10.1080/0263514990170209

Kim, H.-B., Fisher, D. L., & Fraser, B. J. (2000). Classroom environment and teacher interpersonal behaviour in secondary science classes in Korea. *Evaluation &*

Research in Education, 14(1), 3–22. https://doi.org/10.1080/0263514990170209

Kleitman, S., & Gibson, J. (2011). Metacognitive beliefs, self-confidence and primary learning environment of sixth grade students. *Learning and Individual Differences, 21*(6), 728–735. https://doi.org/10.1016/j.lindif.2011.08.003

Kleitman, S., & Stankov, L. (2007). Self-confidence and metacognitive processes. *Learning and Individual Differences, 17*(2), 161–173. https://doi.org/10.1016/j.lindif.2007.03.004

Koul, R. B., Fisher, D. L., & Shaw, T. (2011). An application of the TROFLEI in secondary-school science classes in New Zealand. *Research in Science & Technological Education, 29*(2), 147–167. https://doi.org/10.1080/02635143.2011.573475

Leder, G., & Forgasz, H. (2006). Affect and mathematics education. In A. Gutiérrez & P. Boero (Eds.), *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future* (pp. 403–427). Sense Publishers.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences, 19*(3), 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Lee, S., Fraser, B. J., & Fisher, D. L. (2003). Teacher–student interactions in Korean high school science classrooms. *International Journal of Science and Mathematics Education, 1*(1), 67–85. https://doi.org/10.1023/a:1026191226676

Lewin, K. (1936). *Principles of topological psychology* [Kindle edition]. McGraw.

Li, N., & Koch, I. (2017). *Choose maths gender report: Participation, performance, and attitude towards mathematics*. Australian Mathematical Sciences

Institute. https://amsi.org.au/publications/gender-report-2017-participation-performance-attitudes-towards-mathematics/

Liljedahl, P., & Hannula, M. S. (2016). Research on mathematics-related affect. In Á. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education: The journey continues* (pp. 417–446). SensePublishers.

Lim, S., & Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics, 82*(1), 145–164. https://doi.org/10.1007/s10649-012-9414-x

Lin, S.-H., & Huang, Y.-C. (2014). Development and application of a Chinese version of the short attitudes toward mathematics inventory. *International Journal of Science and Mathematics Education*, *14*(1), 193–216. https://doi.org/10.1007/s10763-014-9563-8

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Logan, K., Crump, B., & Rennie, L. (2006). Measuring the computer classroom environment: Lessons learned from using a new instrument. *An International Journal*, *9*(1), 67–93. https://doi.org/10.1007/s10984-005-9004-2

Lorsbach, A., & Jinks, J. (1999). Self-efficacy theory and learning environment research. *Learning Environments Research, 2*(2), 157–167. https://doi.org/10.1023/A:1009902810926

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Ganley, C. M. (2013). Girls' and boys' mathematics achievement affect and experiences: Findings from ECLS-K. *Journal for Research in Mathematics Education, 44*(4), 634–645. https://doi.org/10.5951/jresematheduc.44.4.0634

Ma, X., & Kishor, N. (1997a). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education, 28*(1), 26–47. https://doi.org/10.2307/749662

Ma, X., & Kishor, N. (1997b). Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review. *Educational Psychology Review, 9*(2), 89–120.

MacLeod, C., & Fraser, B. J. (2010). Development, validation and application of a modified Arabic translation of the What Is Happening In this Class? (WIHIC) questionnaire. *Learning Environments Research, 13*(2), 105–125. https://doi.org/10.1007/s10984-008-9052-5

Majeed, A. A., Darmawan, I. G. N., & Lynch, P. (2013). A confirmatory factor analysis of attitudes toward mathematics inventory (ATMI). *The mathematics educator, 15*(1), 121–135.

Majeed, A. A., Fraser, B. J., & Aldridge, J. M. (2002). Learning environment and its association with student satisfaction among mathematics students in Brunei Darussalam. *Learning Environments Research, 5*(2), 203–226. https://doi.org/10.1023/A:1020382914724

Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*(2), 77–172. https://doi.org/10.1007/bf01322177

Marsh, H. W. (1990b). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement.

Journal of Educational Psychology, 80(3), 366–380. https://doi.org/10.1037/0022-0663.80.3.366

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and ... *Educational Psychologist, 34*(3), 155–167. https://doi.org/10.1207/s15326985ep3403_2

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59–77. https://doi.org/10.1348/000709910x503501

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2017). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263–280. https://doi.org/10.1037/dev0000393

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331–353. https://doi.org/10.1037/edu0000281

Marsh, H. W., Relich, J. D., & Smith, I. D. (1983). Self-concept: The construct validity of interpretations based upon the SDQ. *Journal of Personality and Social Psychology, 45*(1), 173–187. https://doi.org/10.1037/0022-3514.45.1.173

Martin-Dunlop, C., & Fraser, B. J. (2007). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education, 6*(1), 163–190. https://doi.org/10.1007/s10763-007-9070-2

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. Guilford Publications.

International Baccalaureate Organization (2013). *The Middle Years Programme: Providing students with tools that pave the way to higher learning*. https://www.ibo.org/globalassets/publications/recognition/myprecognitionbrochurerevdec2013.pdf

Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Student/teacher relations and attitudes toward mathematics before and after the transition to junior high school. *Child Development, 60*(4), 981–992. https://doi.org/10.2307/1131038

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. The University of Michigan.

Moos, R. H. (1973). Conceptualizations of human environments. *American Psychologist, 28*(8), 652–665. https://doi.org/10.1037/h0035722

Moos, R. H., & Houts, P. S. (1968). Assessment of the social atmospheres of psychiatric wards. *Journal of Abnormal Psychology, 73*(6), 595–604. https://doi.org/10.1037/h0026600

Morony, S., Kleitman, S., Lee, Y. P., & Stankov, L. (2013). Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research, 58*, 79–96. https://doi.org/10.1016/j.ijer.2012.11.002

Mueller, R. J. (2017). Performance assessment. In *Principles of classroom learning and perception* (pp. 242–275). Routledge.

Mulhern, F., & Rae, G. (1998). Development of a shortened form of the Fennema-Sherman Mathematics Attitudes Scale. *Educational and Psychological Measurement, 58*(2), 295–306. https://doi.org/10.1177/0013164498058002012

Murray, H. A. (1938). *Explorations in personality*. Oxford University Press.

National Health and Medical Research Council (Australia) (NHMRC). (2013). *National Statement on Ethical Conduct in Human Research 2007. (Updated May 2013).* National Health and Medical Research Council.

Nielsen, I. L., & Moore, K. A. (2003). Psychometric data on the mathematics self-efficacy scale. *Educational and Psychological Measurement, 63*(1), 128–138. https://doi.org/10.1177/0013164402239321

Niepel, C., Brunner, M., & Preckel, F. (2014). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relations in above average ability secondary school students. *Contemporary Educational Psychology, 39*(4), 301–313. https://doi.org/10.1016/j.cedpsych.2014.07.002

Nix, R. K., Fraser, B. J., & Ledbetter, C. E. (2005). Evaluating an integrated science learning environment using Constructivist Learning Environment Survey. *Learning Environments Research, 8*(2), 109–133. https://doi.org/10.1007/s10984-005-7251-x

Nurmi, A., Hannula, M., Maijala, H., & Pehkonen, E. (2003 July). *On pupils' self-confidence in mathematics: Gender comparisons*. [Conference paper]. The 27th International Group for the Psychology of Mathematics Education Conference, Honolulu, Hawaii, United States.

Nuttal, R. L., Casey, M. B., & Pezaris, E. (2005). Spatial ability as a mediator of gender differences on mathematics tests: A biological - environmental framework. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 121–142). Cambridge University Press.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education, 9*(3), 325–335. https://doi.org/10.1080/1356251042000216642

Office of the Chief Scientist. (2013). *Science, technology, engineering and mathematics in the national interest: A strategic approach*. Commonwealth of Australia. https://www.chiefscientist.gov.au/sites/default/files/STEMstrategy290713FINALweb.pdf

Office of the Chief Scientist. (2014). *Science, technology, engineering and mathematics: Australia's future.* Commonwealth of Australia. https://www.chiefscientist.gov.au/sites/default/files/STEM_AustraliasFuture_Sept2014_Web.pdf

Ogbuehi, P. I., & Fraser, B. J. (2007). Learning environment, attitudes and conceptual development associated with innovative strategies in middle-school mathematics. *Learning Environments Research, 10*(2), 101–114. https://doi.org/10.1007/s10984-007-9026-z

Organization for Economic Co-operation and Development (OECD). (2015). Girls' lack of self-confidence. In *The ABC of gender equality in education: Aptitude, behaviour, confidence* (pp. 63–95). OECD Publishing. http://dx.doi.org/10.1787/9789264229945-6-en.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578. https://doi.org/10.3102/00346543066004543

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative approach* (pp. 294–331). Cambridge University Press.

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: self-efficacy, self-concept, and school achievement. In R. J. Riding & S. Rayner (Eds.), *Self perception* (pp. 239–265). Ablex Publishing.

Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th ed.). Open University Press.

Peer, J. (2011). *Gender, grade-level and stream differences in learning environment and student attitudes in primary science classrooms in Singapore.* [Unpublished doctoral dissertation]. Curtin University.

Peer, J., & Fraser, B. J. (2015). Sex, grade-level and stream differences in learning environment and attitudes to science in Singapore primary schools. *Learning Environments Research, 18*(1), 143–161. https://doi.org/10.1007/s10984-013-9142-x

Pepin, B. (2011). Pupils' attitudes towards mathematics: a comparative study of Norwegian and English secondary students. *ZDM, 43*(4), 535–546. https://doi.org/10.1007/s11858-011-0314-9

Pierce, R., Stacey, K., & Barkatsas, A. (2007). A scale for monitoring students' attitudes to learning mathematics with technology. *Computers & Education, 48*(2), 285–300. https://doi.org/10.1016/j.compedu.2005.01.006

Primi, C., Bacherini, A., Beccari, C., & Donati, M. A. (2020). Assessing math attitude through the Attitude Toward Mathematics Inventory – Short form in introductory statistics course students. *Studies in Educational Evaluation, 64,* 100838. https://doi.org/10.1016/j.stueduc.2020.100838

Punch, K. (1998). *Introduction to social research: Quantitative and qualitative approaches.* Sage Publications.

Quek, C. L., Wong, A. F. L., & Fraser, B. J. (2005a). Student perceptions of chemistry laboratory learning environments, student–teacher interactions and attitudes in secondary school gifted education classes in Singapore. *Research in Science Education, 35*(2–3), 299–321. https://doi.org/10.1007/s11165-005-0093-9

Quek, C. L., Wong, A. F. L., & Fraser, B. J. (2005b). Teacher-student interaction and gifted students' attitudes toward chemistry in laboratory classrooms in Singapore. *Journal of Classroom Interaction, 40*(1), 18–28.

Randhawa, B. S., & Michayluk, J. O. (1975). Learning environment in rural and urban classrooms. *American Educational Research Journal, 12*(3), 265–279. https://doi.org/10.3102/00028312012003265

Reardon, S. F., Kalogrides, D., Fahle, E. M., Podolsky, A., & Zárate, R. C. (2018). The relationship between test item format and gender achievement gaps on math and ELA tests in fourth and eighth grades. *Educational Researcher, 47*(5), 284–294. https://doi.org/10.3102/0013189x18762105

Rennie, L. J. (2010). Gender still matters in Australian schooling. *International Journal of Gender, Science and Technology, 2*(1), 100–111.

Rentoul, A. J., & Fraser, B. J. (1979). Conceptualization of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies, 11*(3), 233–245. https://doi.org/10.1080/0022027790110306

Rentoul, A. J., & Fraser, B. J. (1983). Development of a school-level environment questionnaire. *Journal of Educational Administration, 21*(1), 21–39. https://doi.org/10.1108/eb009866

Reyes, L. (1984). Affective variables and mathematics education. *The Elementary School Journal, 84*(5), 558–581. https://doi.org/10.1086/461384

Reyes, L., & Stanic, G. (1988). Race, sex, socioeconomic status, and mathematics. *Journal for Research in Mathematics Education, 19*(1), 26–43. https://doi.org/10.2307/749109

Rita, R. D., & Martin-Dunlop, C. S. (2011). Perceptions of the learning environment and associations with cognitive achievement among gifted biology students. *Learning Environments Research, 14*(1), 25–38. https://doi.org/10.1007/s10984-011-9080-4

Roberts, K. (2014, 1 May). *Engaging more women and girls in mathematics and STEM fields: The international evidence*. Australian Mathematical Sciences

Institute. https://amsi.org.au/wp-content/uploads/2014/08/RobertsGenderSTEMreport2014.pdf

Schöber, C., Schütte, K., Köller, O., McElvany, N., & Gebauer, M. M. (2018). Reciprocal effects between self-efficacy and achievement in mathematics and reading. *Learning and Individual Differences, 63*, 1–11. https://doi.org/10.1016/j.lindif.2018.01.008

Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement, 51*(1), 67–78. https://doi.org/10.1177/0013164491511005

Sebela, M., Fraser, B. J., & Aldridge, J. M. (2003). *Teacher action research and constructivist classroom environments in South Africa* [Conference paper]. Third International Conference on Science, Mathematics and Technology Education, Rhodes University, Makhanda, South Africa.

Seidman, E., Allen, L., Aber, J. L., Mitchell, C., & Feinman, J. (1994). The impact of school transitions in early adolescence on the self-system and perceived social context of poor urban youth. *Child Development, 65*(2), 507–522. https://doi.org/10.2307/1131399

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research, 46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Smith, E. (2010). Women into science and engineering? Gendered participation in higher education STEM subjects. *British Educational Research Journal, 37*(6), 993–1014. https://doi.org/10.1080/01411926.2010.515019

Song, I.-S., & Hattie, J. (1984). Home environment, self-concept, and academic achievement: A causal modelling approach. *Journal of Educational Psychology, 76*(6), 1269–1281. https://doi.org/10.1037/0022-0663.76.6.1269

Speering, W., & Rennie, L. J. (1996). Students' perceptions about science: The impact of transition from primary to secondary school. *Research in Science Education, 26*(3), 283–298. https://doi.org/10.1007/bf02356940

Stankov, L. (1999). Mining on the 'No man's land' between intelligence and personality. In P. Ackerman, P. Kyllonen, & R. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 315–338). American Psychological Association.

Stankov, L., Kleitman, S., & Jackson, S. A. (2015). Measures of the trait of confidence. In D. H. Saklofske & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (pp. 158–189). Academic Press.

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science, 29*(4), 581–593. https://doi.org/10.1177/0956797617741719

Stones, E. (1972). *An introduction to educational psychology*. Methuen & Co.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson

Tapia, M. (1996 6 November). *The attitudes toward mathematics instrument* [Conference paper]. The Annual Meeting of the mid-South Educational Research Association, Tuscaloosa, Alabama, United States.

Tapia, M., & Marsh, G. E., II. (2000 16 November). *Attitudes toward mathematics instrument: An investigation with middle school students* [Conference paper]. The Annual meeting of the mid-South Educational Research Association, Bowling Green, Kentucky, United States.

Tapia, M., & Marsh, G. E., II. (2002 6 November). *Confirmatory factor analysis of the attitudes toward mathematics inventory* [Conference paper]. The Annual Meeting of the mid-South Educational Research Association, Chattanooga, Tennessee, United States.

Tapia, M., & Marsh, G. E., II. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly, 8*(2), 16–21.

Tapia, M., & Marsh, G. E., II. (2005). Attitudes toward mathematics inventory redux. *Academic Exchange Quarterly, 9*(3), 272–278.

Tartre, L. A., & Fennema, E. (1995). Mathematics achievement and gender: A longitudinal study of selected cognitive and affective variables [Grades 6–12]. *Educational Studies in Mathematics, 28*(3), 199–217. https://doi.org/10.1007/bf01274173

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*(4), 293–302. https://doi.org/10.1016/s0883-0355(97)90011-2

Telli, S., Cakiroglu, J., & Brok, P. d. (2006). Turkish secondary education students' perceptions of their classroom learning environment and their attitude towards biology. In D. L. Fisher & M. S. Khine (Eds.), *Contemporary approaches to research on learning environments: Worldviews* (pp. 517–542). World Scientific Publishing Co.

Thomson, S., Hillman, K., Wernett, N., Schmid, M., Buckley, S., & Munene, A. (2012). *Highlights from TIMMS & PIRLS 2011 from Australia's perspective*. Australian Council for Educational Research. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1001&context=timss_pirls_2011

Thomson, S., Wernett, N., O'Grady, E., & Rodrigues, S. (2017). *TIMSS 2015: Reporting Australia's results*. Australian Council for Educational Research.

Thomson, S., Wernett, N., Underwood, C., & Nicholas, M. (2008). *TIMSS 07: Taking a closer look at mathematics and science in Australia*. Australian Council of Education Research.

Tobin, K., & Fraser, B. J. (1998). Qualitative and quantitative landscapes of classroom learning environments. In B. J. Fraser & K. G. Tobin (Eds.), *The international handbook of science education* (pp. 623–640). Kluwer Academic Publishers.

Tran, V. D. (2012). Predicting the attitudes and self-esteem of the grade 9th lower secondary school students towards mathematics from their perceptions of the classroom learning environment. *World Journal of Education, 2*(4), 34–44. https://doi.org/10.5430/wje.v2n4p34

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology, 90*(2), 334–349. https://doi.org/10.1037/0022-3514.90.2.334

Trickett, E. J., & Moos, R. H. (1973). Social environment of junior high and high school classrooms. *Journal of Educational Psychology, 65*(1), 93–102. https://doi.org/10.1037/h0034823

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2012). *World atlas of gender equity in education*. UNESCO.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Validakis, V. (2015, 29 April). *BHP Billiton wants more girls to study maths*. Australian Mining. https://www.australianmining.com.au/news/bhp-billiton-wants-more-girls-to-study-maths/

Vanayan, M., White, N., Yuen, P., & Teper, M. (1997). Beliefs and attitudes towards mathematics among third- and fifth-grade students: A descriptive study. *School Science and Mathematics, 97*(7), 345–351. https://doi.org/10.1111/j.1949-8594.1997.tb17375.x

Velayutham, S., & Aldridge, J. M. (2013). Influence of psychosocial classroom environment on students' motivation and self-regulation in science learning: A structural equation modeling approach. *Research in Science Education, 43*(2), 507–527. https://doi.org/10.1007/s11165-011-9273-y

Velayutham, S., Aldridge, J. M., & Fraser, B. J. (2011). Development and validation of an instrument to measure students' motivation and self-regulation in science learning. *International Journal of Science Education*, *33*(15), 2159–2179. https://doi.org/10.1080/09500693.2010.541529

Waddington, J. (2019). Developing primary school students' foreign language learner self-concept. *System*, *82*, 39–49. https://doi.org/10.1016/j.system.2019.02.012

Walberg, H. J. (1969). Predicting class learning: An approach to the class as a social system. *American Educational Research Journal*, *6*(4), 529–542. https://doi.org/10.3102/00028312006004529

Walberg, H. J. (1976). Psychology of learning environments: Behavioral, structural, or perceptual? *Review of Research in Education, 4*, 142–178. https://doi.org/10.2307/1167115

Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership*, *41*(8), 19–27.

Walberg, H. J., & Anderson, G. J. (1968a). Classroom climate and individual learning. *Journal of Educational Psychology*, *59*(6), 414–419. https://doi.org/10.1037/h0026490

Walberg, H. J., & Anderson, G. J. (1968b). *Learning environment inventory (1967–68 edition).*Educational Testing Service.

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science, 24*(5), 770–775. https://doi.org/10.1177/0956797612458937

Watt, H. M. G. (2005). Explaining gendered math enrollments for NSW Australian secondary school students. *New Directions for Child and Adolescent Development, 2005*(110), 15–29. https://doi.org/10.1002/cd.147

Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educational Research and Evaluation, 12*(4), 305–322. https://doi.org/10.1080/13803610600765562

Watt, H. M. G. (2007). A trickle from the pipeline: Why girls under participate in maths. *Professional Educator, 6*(3), 36–41.

Watt, H. M. G., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls: A motivational analysis of high school enrolments in Australia and the USA. *Equal Opportunities International, 25*(8), 642–659. https://doi.org/10.1108/02610150610719119

Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., & Harackiewicz, J. M. (2017). Mathematics – a critical filter for STEM-related career choices? A longitudinal examination among Australian and U.S. adolescents. *Sex Roles, 77*(3–4), 254–271. https://doi.org/10.1007/s11199-016-0711-1

Way, N., Reddy, R., & Rhodes, J. (2007). Students' perceptions of school climate during the middle school years: Associations with trajectories of psychological and behavioral adjustment. *American Journal of Community Psychology, 40*(3–4), 194–213. https://doi.org/10.1007/s10464-007-9143-y

Wigfield, A., Eccles, J., Mac Iver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology, 27*(4), 552–565. https://doi.org/10.1037/0012-1649.27.4.552

Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education, 38*(3), 321–341. https://doi.org/10.1007/s11165-007-9052-y

Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., & Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: A Dynamic test of the big-fish–little-pond effect. *Journal of Educational Psychology, 104*(3), 793–805. https://doi.org/10.1037/a0027732

Wubbels, T. (1993). *Teacher-student relationships in science and mathematics classes.* Curtin University of Technology.

Wubbels, T., & Levy, J. (Eds). (1993). *Do you know what you look like?: Interpersonal relationships in education*. Routledge Falmer.

Yang, X. (2015). Rural junior secondary school students' perceptions of classroom learning environments and their attitude and achievement in mathematics in West China. *Learning Environments Research, 18*(2), 249–266. https://doi.org/10.1007/s10984-015-9184-3

Yarrow, A., Millwater, J., & Fraser, B. J. (1997). Improving university and primary school classroom environments through preservice teachers' action research. *International Journal of Practical Experiences in Professional Education, 1*(1), 68–93.

Yee, L. S. (2010 July). *Mathematics attitudes and achievement of junior college students in Singapore.* [Conference paper]. Shaping the Future of Mathematics Education: The 33rd Annual Conference of the Mathematics Education Research Group of Australasia, Fremantle, Western Australia, Australia.

Yeung, A. S. (2011). Student self-concept and effort: gender and grade differences. *Educational Psychology*, *31*(6), 749–772. https://doi.org/10.1080/01443410.2011.608487

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal*, *37*(1), 215–246. https://doi.org/10.3102/00028312037001215

Zevenbergen, R. (2003). Grouping by ability: A self-fulfilling prophecy? *The Australian Mathematics Teacher*, *59*(4), 2–7. https://doi.org/10.3102/00028312037001215

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

Appendix A

COPY OF M-ATMI

Confidence measures the students' confidence and their self-concept of their performance in mathematics.

1. I am confident when it comes to mathematics.
2. I expect to do well in mathematics.
3. I learn mathematics easily.
4. I am confident that I can learn advanced mathematics.
5. I am comfortable answering questions in mathematics lessons.
6. I am good at solving mathematics problems.

Value measures the students' beliefs about the worth of mathematics in daily life now and in the future.

7. Mathematics is a worthwhile subject.
8. Mathematics teaches a person to think.
9. Mathematics is one of the most important subjects for people to study.
10. I use mathematics in my daily life.
11. Studying mathematics helps me to problem solve in other areas.
12. Mathematics will be helpful to me in the future.

Enjoyment measures how much students enjoy doing mathematics.

13. I look forward to mathematics lessons.
14. Mathematics lessons are fun.
15. Mathematics lessons interest me.
16. Mathematics is one of my favourite school subjects.
17. There should be more mathematics lessons.
18. I enjoy the activities we do in mathematics lessons.

Motivation measures the students' interest in mathematics.

19. I want to learn new mathematics content.
20. I want to master new mathematics skills.
21. It is important for me to learn the mathematics content that is taught.
22. It is important for me that I improve my mathematics skills.
23. Understanding mathematics is important to me.

Appendix B

COPY OF WIHIC

Student Friendship/Cohesiveness

1. I make friendships among students in this class.
2. I know other students in this class.
3. I am friendly to members of this class.
4. Members of the class are my friends.
5. I work well with other class members.
6. Students in this class like me.

Teacher Support

7. The teacher considers my feelings.
8. The teacher helps me when I have trouble with my work.
9. The teacher talks with me.
10. The teacher takes an interest in my progress.
11. The teacher moves about the class to talk to me.
12. The teacher's questions help me to understand.

Equity

13. The teacher gives as much attention to my questions as to other students' questions.
14. I get the same amount of help from the teacher as do other students.
15. I have the same amount of say in this class as other students.
16. I receive the same encouragement from the teacher as other students do.
17. I get the same opportunity to contribute to class discussions as other students.
18. I get the same opportunity to answer questions as other students.

Involvement

19. I discuss ideas in class.
20. I give my opinions during class discussions.
21. The teacher asks me questions.
22. My ideas and suggestions are used during class discussions.
23. I explain my ideas to other students.
24. I am asked to explain how I solve problems.

Task Focus/Orientation

25. Getting a certain amount of work done is important to me.
26. I am ready to start this class on time.
27. I set my own goals for this class.
28. I pay attention during this class.
29. I try to understand the work in this class.
30. I know how much work I have to do.

Cooperation

31. We work in groups (or pairs) in this class.
32. When I work in groups in this class, there is teamwork.
33. I work with other students on assignments in this class.
34. I cooperate with other students on class activities.
35. I share my books and resources with other students when doing class work.
36. I learn from other students in this class.

Clarity of Assessment Criteria

37. I am aware of which activities and tasks are used to assess my performance.
38. I know what types of information are needed to complete an assessment task.
39. The instructions for assessment are clear to me.
40. I know how to complete assessment tasks successfully.
41. I understand how the teacher judges my work.
42. I know how to complete different assessment tasks.

Appendix C

INSTRUCTIONS FOR COMPLETING THE QUESTIONNAIRES.

Thank you for participating in my survey. Your feedback is important.

This questionnaire has three parts.

Part 1 asks you a couple of questions about yourself.

Part 2 asks you to describe your mathematics class. The one which you are in right now(Actual) and the one you would like to be in (Preferred).

Part 3 asks you about your attitudes to mathematics.

This is not a test.

There are no right or wrong answers.

Your opinion is what is wanted.

Your answers will be confidential.

You must answer every question.

Part 1

Which mathematics class are you in?

1. Year level.
2. The name of your mathematics teacher.
3. Which topic are you studying in your current mathematics class?

Now you are ready to start Part 2

In Part 2 there are 42 statements about practices which could take place in your mathematics class. You will be asked how often each practice takes place. Choose one answer (Almost Never, Seldom, Sometime, Often, Almost Always) for each row by clicking in the circle.

The first row is describing the mathematics class you are in at the moment (Actual). This about how well each statement describes what your mathematics class is like for you.

The second row is describing the mathematics class you would like to be in (Preferred). Think about how well each statement describes the mathematics class you would like to be in.

For example:

The teacher helps me when I have trouble with my work.

If you think this happens sometimes in your current mathematics class, click the circle under the word Sometimes in the Actual row.

If you would like this happen more often in your mathematics class, choose Often or Almost Always from the Preferred row.

If you want to change an answer just click in a different circle. You can go back to look at previous questions or to change your answers, but you will not be able to go forward until you have recorded an answer for each question on the screen.

There are no right or wrong answers. Your opinion is what is wanted.

You have finished Part 2

In Part 3 there are twenty three statements about your Attitude to Mathematics.

For each question choose one of the five answers: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree.

For example:

Mathematics is a worthwhile subject.

If you agree with this statement click the circle under the word Agree.

If you want to change an answer just click in a different circle. You can go back to change previous answers, but you will not be able to go forward until you have answered each question on the screen.

Remember there are no right and wrong answers.

You have finished the questionnaire.

Thank you for participating.

Appendix D

COPY OF ETHICS APPROVAL

Curtin University

Memorandum

To	Sharon Quinn, SMEC
From	Mun Yin Cheong, Form C Ethics Co-ordinator Faculty of Science and Engineering
Subject	Protocol Approval SMEC-39-14
Date	7 July 2014
Copy	Jill Aldridge, SMEC

Office of Research and Development

Human Research Ethics Committee

Telephone 9266 2784
Facsimile 9266 3793
Email hrec@curtin.edu.au

Thank you for your "Form C Application for Approval of Research with Low Risk (Ethical Requirements)" for the project titled *"Investigating relationships between girls' confidence in mathematics and their perceptions of their learning environment"*. On behalf of the Human Research Ethics Committee, I am authorised to inform you that the project is approved.

Approval of this project is for a period of 4 years **7th July 2014** to **6th July 2018**.

Your approval has the following conditions:

(i) Annual progress reports on the project must be submitted to the Ethics Office.

(ii) It is your responsibility, as the researcher, to meet the conditions outlined above and to retain the necessary records demonstrating that these have been completed.

The approval number for your project is **SMEC-39-14**. Please quote this number in any future correspondence. If at any time during the approval term changes/amendments occur, or if a serious or unexpected adverse event occurs, please advise me immediately.

Regards,

Mun Yin

MUN YIN CHEONG
Form C Ethics Co-ordinator
Faculty of Science and Engineering

Please Note: The following standard statement must be included in the information sheet to participants:
This study has been approved under Curtin University's process for lower-risk Studies (Approval Number xxxx). This process complies with the National Statement on Ethical Conduct in Human Research (Chapter 5.1.7 and Chapters 5.1.18-5.1.21).
For further information on this study contact the researchers named above or the Curtin University Human Research Ethics Committee. c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth 6845 or by telephoning 9266 9223 or by emailing hrec@curtin.edu.au.

Standard conditions of ethics approval

These standard conditions apply to all research approved by the Curtin University Human Research Ethics Committee. It is the responsibility of each researcher named on the application to ensure these conditions are met.

1. **Compliance.** Conduct your research in accordance with the application as it has been approved and keep appropriate records.
    a. **Monitoring** – Assist the Committee to monitor the conduct of the approved research by completing promptly and returning all project review forms that are sent to you.
    b. **Annual report** - Submit an annual report on or before the anniversary of the approval.
    c. **Extensions** - If you are likely to need more time to conduct your research than is already approved, complete a new application six weeks before the current approval expires.
    d. **Changes to protocol** – Any changes to the protocol are to be approved by the Committee before being implemented.
    e. **Changes to researcher details** - Advise the Committee of any changes in the contact details of the researchers involved in the approved study.
    f. **Discontinuation** - You must inform the Committee, giving reasons, if the research is not conducted or is discontinued before the expected completion date.
    g. **Closure** – Submit a final report when the research is completed. Include details of when data will be destroyed, and how, or if any future use is planned for the data.
    h. **Candidacy** - If you are a Higher Degree by Research student, data collection must not begin before your Application for Candidacy is approved by your Faculty Graduate Studies Committee.
2. **Adverse events.** Consider what might constitute an adverse event and what actions may be needed if an adverse event occurs. Follow the procedures for reporting and addressing adverse events (http://research.curtin.edu.au/guides/adverse.cfm). Where appropriate, provide an adverse events protocol. The following are examples of adverse events:
    a. Complaints
    b. Harm to participants. This includes physical, emotional, psychological, economic, legal, social and cultural harm (NS Section 2)
    c. Loss of data or breaches of data security
    d. Legal challenges to the research
3. **Data management plan.** Have a Data Management Plan consistent with the University's recordkeeping policy. This will include such things as how the data are to be stored, for how long, and who has authorised access.
4. **Publication.** Where practicable, ensure the results of the research are made available to participants in a way that is timely and clear (NS 1.5). Unless prohibited from doing so by contractual obligations, ensure the results of the research are published in a manner that will allow public scrutiny (NS 1.3, d). Inform the Committee of any constraints on publication.
5. **Police checks and other clearances.** All necessary clearances, such as Working with Children Checks, first aid certificates and vaccination certificates, must be obtained before entering a site to conduct research.
6. **Participant information.** All information for participants must be approved by the HREC before being given to the participants or made available to the public.
    a. **University logo.** All participant information and consent forms must contain the Curtin University logo and University contact details for the researchers. Private contact details should not be used.
    b. **Standard statement.** All participant information forms must contain the HREC standard statement.
    This study has been approved under Curtin University's process for lower-risk Studies (Approval Number xxxx). This process complies with the National Statement on Ethical Conduct in Human Research (Chapter 5.1.7 and Chapters 5.1.18-5.1.21).
    For further information on this study contact the researchers named above or the Curtin University Human Research Ethics Committee. c/- Office of Research and Development, Curtin University, GPO Box U1987, Perth 6845 or by telephoning 9266 9223 or by emailing hrec@curtin.edu.au.
    c. **Plain language.** All participant information must be in plain language that will be easily understood by the participants.

Please direct all communication through the your Form C Ethics Co-ordinator.

* Form C Coordinator cannot not approve amendment request, these must be approved by a Form C reviewer.

Appendix E

LETTER TO PRINCIPALS

My address

1st July, 2015

The Principal
School address

Dear…..,

My name is Sharon Quinn, after twenty seven years teaching mathematics in Adelaide Catholic schools I am now a PhD student at Curtin University. I am conducting a research project on girls and mathematics which will form the basis for my Doctor of Philosophy: Science and Mathematics Education. The aim of the study is to examine the relationships between elements of the mathematics learning environment, girls' confidence in their ability to do mathematics and their achievement. I hope that identifying these relationships will increase our understanding of what influences female students' confidence in their ability to do mathematics. I am writing to request the participation of your Year 8 to 10 students in my research.

The study will involve students in completing a fifteen minute on-line questionnaire three times over two terms. The questionnaire will include questions about students' current mathematics learning environment and their attitudes towards mathematics. The students will not be identifiable to maintain confidentiality and the only people with access to the data will be me and my supervisor. More details of the study are included in the Teacher and Parent Information sheets.

I will strive to minimise the impact of my study on the normal mathematics programme and teacher workloads. However, I will need teachers to hand out and collect student permission forms and to provide a list of student results for each unit of work assessed during the study.

Included is a copy of my approval letter from the Ethics coordinator of the Faculty of Science and Engineering (approval number SMEC – 39 – 14). In addition, I have included a copy of the questions in the questionnaire (this will be administered using

SurveyMonkey), the Parent Information Sheet and Consent Form, the Teacher Information Sheet and a copy of my current Certificate of Teacher Registration.

I would welcome the opportunity to meet with you to discuss my research and to provide any further information you may require. My contact details are: mobile: 0417 894 240,
home: 8344 5331, email: sharon.quinn@adelaide.on.net.

Yours faithfully

Sharon Quinn

Appendix F

APPROVAL FROM SCHOOL 1

Dear Sharon

Apologies for the delay and yes, as you say, the end of term is hectic as usual.

I have checked with Fiona and Liz Min who as you know are the Middle and Senior Years Maths Coordinators and they are happy to work with you on the research. I am happy to support your research in this area Sharon. I suggest you make contact directly with Fiona and Liz and work with them and keep me in the loop as to the progress. I will pass on your package of information to them to follow up.

I wish you luck in your research and look forward to the results.

Best wishes

[redacted]

[redacted]

Appendix G

APPROVAL FROM SCHOOL 2

Hi Sharon

Just letting you know that we are willing to help you out!

What is the next step?

Regards

[redacted]

Mathematics Learning Area Leader

Appendix H

STUDENT INFORMATION SHEET AND CONSENT FORM

Information Sheet

Dear Parents,

My name is Sharon Quinn, I am a student at Curtin University. I am conducting a study on girls and mathematics which will form the basis for a Doctor of Philosophy: Science and Mathematics Education. The title of the study is: Investigating relationships between girls' confidence in mathematics and their perceptions of their learning environment. The aim of the study is to examine the relationships between elements of the mathematics learning environment, girls' confidence in their ability to do mathematics and their achievement in mathematics.

Participants will complete a fifteen minute on-line questionnaire which will be administered during a mathematics lesson in term 4. The students will be asked about the learning environment in their mathematics class and their attitudes towards mathematics. The teachers will provide me with students' grade.

All aspects of the study will be strictly confidential and apart from me only my supervisor will have access to information about participants. The data will be statistically analysed to explore the relationships between elements of the mathematics learning environment, student confidence and achievement. I will provide a summary of the results to the teachers and I will be writing and submitting a thesis for publication. Individual participants will not be identified in either the summary or the thesis. The data will be stored on a password protected computer at Curtin University and will be held for 7 years after which they will be destroyed.

Participation in this study will be beneficial in investigating the relationships between student perceptions of the mathematics learning environment and their confidence in their ability to do mathematics. Identifying these relationships will increase our understanding of what influences students' confidence in their ability to do mathematics.

Your daughter's participation in the study is completely voluntary and she may withdraw from the study at any time, or you may withdraw her from the study at any time. If this occurs all records of your daughter's participation in the study will be destroyed.

If you have any questions or would like more information please contact me.

Contact details

Researcher: Sharon Quinn, Mob: 0439 039 319, Email: s.quinn@student.curtin.edu.au

Supervisor: Associate Professor Jill Aldridge, Email: J.Aldridge@curtin.edu.au

Consent Form

I (print name) …………………………………………………………..GIVE / DO NOT

GIVE consent to the participation of my daughter (print name)

………………………………………………… in the research project described

below.

Title of project: Investigating relationships between girls' confidence in mathematics and their perceptions of their learning environment.

Researcher: Sharon Quinn
Supervisor: Jill Aldridge, Associate Professor: Science and Mathematics Education, Curtin University

In giving consent I acknowledge that:

1. The procedures required for the project and the time involved have been explained to me and any questions I have about the project have been answered to my satisfaction
2. I have read the Parent Information Sheet and have been given the opportunity to discuss the information and my daughter's involvement in the project with the researcher.
3. I have discussed participation in the project with my daughter and my daughter assents to her participation in the project
4. I understand that my daughter's participation in the project is voluntary, a decision not to participate will in no way affect her academic standing or relationship with the school and she is free to withdraw her participation at any time.
5. I understand that my daughter's participation is strictly confidential and that no information about my daughter will be used in any way that reveals my daughter's identity

Signed (Parent):____

Signed (Student): ____

Date: ____

Please return to your Mathematics teacher by Friday 25th September 2015

Appendix I

TEACHER INFORMATION SHEET

Dear Teachers,

My name is Sharon Quinn, I am a student at Curtin University. I am conducting a study on girls and mathematics which will form the basis for a Doctor of Philosophy: Science and Mathematics Education. The title of the study is: Investigating relationships between girls' confidence in mathematics and their perceptions of their learning environment. The aim of the study is to examine the relationships between elements of the mathematics learning environment, girls' confidence in their ability to do mathematics and achievement.

I will be asking students to complete a fifteen minute on-line questionnaire three times over two terms in 2015. The questionnaire will be administered at the end of three different units of work. It will be made up of seven scales from the Constructivist-Orientated Learning Environment Survey (COLES) and four attitude scales. The learning environment scales to be measured are: Student cohesiveness (the extent to which students are supportive of each other); Teacher support (the extent to which the teacher relates to the students); Involvement (the extent to which students participate in the class); Task orientation (the extent to which students stay on task in lessons); Cooperation (the extent to which students cooperate rather than compete); Equity (the extent to which students are treated equally by the teacher) and Assessment Criteria (the extent to which assessment criteria are explicit so that the basis for judgement is clear and public). These scales are made up of six statements each of which is responded to using a five-point frequency scale with the alternatives: Almost Never, Seldom, Sometimes, Often and Almost Always. The four attitude scales are: Self-confidence (which measures students' confidence and self-concept of their performance in mathematics); Value (which measures students' beliefs about the worth of mathematics in their daily lives); Enjoyment (which measures how much students enjoy doing mathematics) and Motivation (which measures students' interest in mathematics). Each of these scales has six statements which are responded to using a five-point Likert scale of strongly disagree, disagree, neutral, agree and strongly agree.

I will be asking you:

- to distribute and collect the Parent/student information sheet and consent form;
- to release those students, whose parents consent to their participation, from class for approximately 20 minutes so that I can administer the questionnaires;
- to provide me with the students' marks for the three units of work completed during the study.

All aspects of the study will be strictly confidential and apart from me only my supervisor will have access to information about participants or individual classes. The data will be statistically analysed to explore the relationships between elements of the mathematics learning environment, student confidence and achievement. I will provide a summary of the results of each class to the teacher of that class. I will also be submitting a thesis for publication. Individual participants will not be identified in either the summary or the thesis. The data will be stored on a password protected computer at Curtin University and will be held for 7 years after which they will be destroyed.

Participation in this study will be beneficial in investigating the relationships between student perceptions of the mathematics learning environment and their confidence in their ability to do mathematics. Identifying these relationships will increase our understanding of what influences students' confidence in their ability to do mathematics.

Student participation in the study is completely voluntary and students may withdraw from the study at any time. If this occurs all records of their participation in the study will be destroyed.

If you have any questions or would like more information please contact me.

Contact details

Researcher: Sharon Quinn, Mob: 0439 039 319, Email: s.quinn@student.curtin.edu.au

Supervisor: Associate Professor Jill Aldridge, Email: J.Aldridge@curtin.edu.au

Appendix J

SUPERVISOR NOTES SCHOOL 2

Supervising the Questionnaire

Students will receive an email from s.quinn@student.curtin.edu.au the subject will be Curtin University Mathematics Research.

If the email is not in their Inbox they should look in their Junk Box and then drag it to their Inbox.

When the students open the email they will find a link to the questionnaire which is in SurveyMonkey. Clicking on the link will take them straight to the questionnaire.

The email links are unique to each student so there is no point them emailing it to another student who may have deleted their email, the link will not work.

Before they start:

> Please explain to the students the difference between actual and preferred:
>
> Actual is thinking about their current mathematics class as they perceive it;
>
> Preferred is thinking about how they would like the class to be if they would like it to be different.

And remind students that

- It is not a test; there are no right or wrong answers it their opinions that are required
- There answers will be confidential i.e. you will not see their answers
- If they have questions they should ask you, not each other - it is okay to answer their questions in terms of explaining the meaning of questions or words. If you need to do this can you please make a note of which questions or words required explanation.

The questionnaire will be open until Wednesday 11th November. If students are away on the day you do it in class can you please encourage them to complete it in another lesson or in their own time? After 9th November I will send a follow-up email to those students who have not completed the questionnaire and ask them to complete it in their own time.

Thanks for your support with this project.

The following instructions appear at the beginning of the questionnaire and the beginning of Parts 1 & 2:

Thank you for participating in my survey.

Your feedback is important to understanding how you feel about your mathematics lessons.

This is not a test.

There are no right or wrong answers.

Your opinion is what is wanted.

In Part 1 there are 42 statements about practices which could take place in your mathematics class. You will be asked how often each practice takes place.

Choose one answer (Almost Never, Seldom, Sometimes, Often or Almost Always) for each row by clicking in the circle.

The first row is describing the mathematics class you are in at the moment (Actual). Think about how well each statement describes what your mathematics class is like for you.

The second row is describing the mathematics class you would like to be in (Preferred). Think about how well each statement describes the mathematics class you would like to be in.

In the Part 2 there are twenty three statements about your Attitude to Mathematics.

For each question choose one of the five answers: Strongly Disagree, Disagree, Neutral, Agree or Strongly Agree.

For example: Mathematics is a worthwhile subject.

If you agree with this statement click the circle under the word Agree.

If you want to change an answer just click in a different circle. You can go back to change previous answers, but you will not be able to go forward until you have answered each question on the screen.

Remember

There are no right and wrong answers.

Your opinion is what is wanted.

.