crpp Centre for Research in Pedagogy and Practice

FINAL RESEARCH REPORT
for Project No. CRP 24/03 FLH

February 2008

INTEGRATING NEW ASSESSMENT STRATEGIES INTO MATHEMATICS CLASSROOMS

An Exploratory Study in Singapore Primary and Secondary Schools

Fan Liang Huo, Quek Khiok Seng, Koay Phong Lee, Juliana Donna Ng Chye Huat, Lionel Pereira-Mendoza, Yeo Shu Mei, Tan-Foo Kum Fong, Teo Soh Wah and Zhu Yan

Contents

Please contact the authors for the appendices.

Chapter 1 Introduction

1.1 Background of Research

Since the early 1980s, educational researchers and practitioners in many countries have increasingly realized the inadequacies of traditional assessment methods, which are typically based on paper-and-pencil or written tests. Such written tests are usually time limited—in a block of time, venue limited—within classrooms, and tool limited—paper-and-pencil. Because of those limitations, they often heavily focus on students' memorization of knowledge and familiarity with routine procedures and skills; they cannot effectively assess students for a full range of educational goals and instructional objectives, including students' conceptual understanding, higher-order thinking and creativity, problem solving ability, and communication skills. In addition, such written-test-based assessments have generally excluded students' affect in learning mathematics (for a more detailed discussion of the disadvantages of using such traditional assessment, see Fan, 2006).

To overcome the inadequacies of the traditional written-test-based assessment and to better reflect the new desired instructional goals and shifted values in education, educational researchers and reformers have proposed and implemented a wide range of new assessment strategies, or the so-called "alternative assessment methods", in mathematics classrooms for the last two decades. They included the assessment methods based on project work, performance tasks, journal writing, oral presentation, portfolio, interview and classroom observation, and so on (e.g., see Adam, 1998; Berenson & Carter, 1995; Clarke, 1997; Haines & Izard, 1994; Kulm, 1994; Raymond, 1994; Richardson, 1988; Stacey, 1987; Stempien & Borasi, 1985; Stenmark, Beck, & Asturias, 1994; Zehavi, Bruckheimer, & Ben-Zvi, 1988).

In Singapore, new strategies for assessing students' learning in mathematics classrooms have also received attention from educational policy makers, administrators, researchers, and practitioners, particularly since the mid 1990s (e.g., Fan & Yeo, 2000; Yeo, 2001; Ministry of Education, 2004a, 2004b; Fan, 2002; Seto, 2002). However, in general, there has been a lack of research in this area, particularly relatively large-scale research in the Singapore educational context.

The report presented herein is about a two-year project in the mathematics assessment domain. The project is officially entitled "Integrating New Assessment Strategies into Mathematics Classrooms", but also called "Mathematics Assessment Project" (MAP), for short.

The MAP project was officially launched in December 2003. As a subject-specific project, it was approved and funded by/through the Center for Research in Pedagogy and Practice (CRPP) of National Institute of Education, Nanyang Technological University. The CRPP was established in early 2003 through a large grant from the Singapore Ministry of Education (MOE), as part of its new initiatives and support for educational reform and research.

1.2 Research Purposes and Questions

The overall aim of the MAP project is to provide research-based evidence and practical suggestions for promoting the effective use of new assessment strategies in Singapore mathematics classrooms.

Specifically, the MAP project has the following major objectives:

(a) Investigate the effects of using new assessment strategies in mathematics teaching and learning on students' achievements both in cognitive and affective domains in our local school settings.
(b) Identify practical issues, difficulties, and challenges teachers face in using those new ways of assessment in Singapore's educational context.
(c) Explore the implications of the results from Part A and Part B above and therefore offer practical guidance and advice for teachers to integrate new assessment strategies more effectively into their daily teaching.
(d) Produce research-based evidence and therefore make suggestions about how the researchers and teacher educators (NIE), school administrators (Schools), and educational policy makers (MOE) can help teachers develop their knowledge, skills, and willingness in the use of alternative assessment in classrooms.

It is hoped that the research can also make meaningful contribution to both local and international community of mathematics education concerning the understanding of the values and issues around alternative assessment in mathematics learning and teaching.

In general, there are three broad research questions in this project:

(1) What are the influences of "new assessment strategies" on students' learning of mathematics in their cognitive domain?
(2) What are the influences of "new assessment strategies" on students' learning of mathematics in their affective domain?
(3) How can "new assessment strategies" be effectively integrated into mathematics classrooms?

In this study, by new assessment strategies we mainly mean the process of assessing students' learning of mathematics through their work on project tasks, performance tasks, communication tasks and student self-assessment tasks (see more details about these terms below).

1.3 Report Organization

Since it was launched, the MAP project team and its sub-study teams have organized a number of workshops, presented many papers in local and international conferences, and published book chapters and journal articles, based on the progress made along the way. (See Appendix 1.1 for more information)

What presented here is the final report, providing an overall description of the project, including mainly its conceptualization, methodology, results and findings, and conclusions and implications.

This report is organized into 11 chapters.

Chapter 2 explains the *general* conceptual and methodological matters concerning the research undertaken in this project, including the conception about assessment and alternative assessment, research methods and procedural at the project level .

Chapters 3 to 6 report the sub-studies in terms of each of the four new assessment strategies, that is, performance tasks, project work, communication tasks, and student self-assessment, respectively, in primary school mathematics classrooms.

Chapters 7 to 10 present the sub-studies into the use of the new assessment strategies in secondary school mathematics classrooms, respectively.

In Chapters 3 to 10, we will also describe *specific* conceptual and methodological issues about each particular new assessment strategy in primary and secondary school levels that the chapter addresses. These chapters are developed and based on the initial reports we obtained from the sub-project team members.

Chapter 11 provides a summary of the findings and conclusions drawn from the results obtained from the project. It also offers relevant suggestions and recommendations for policy makers, school administrators, mathematics teachers, and researchers concerning the use of new assessment strategies in school classrooms.

In addition to some research instruments, a large number of the sample invention tasks are provided in the appendices for illustration purposes.

Chapter 2 Conception and Methodology of Research

2.1 Conceptions about Assessment

Traditionally, assessment in education was largely equated with paper-and-pencil (i.e., written) tests (e.g., see Bayles, 1950). In such written tests, what is assessed is students' academic achievement, mainly their knowledge on topics well defined in the curriculum and skills on solving routine and conventional problems, and the most important purpose is to grade and report students' learning results (Fan, 2005).

In the Tenth International Congress on Mathematical Education (ICME-10) held in Denmark, 2004, two terms -"internal assessment" and "external assessment"- were used to distinguish assessment themes. As Stephens explained, internal assessment basically refers to classroom assessment, which is "in the hands of the teacher with the constraints imposed upon teachers and schools by externally sanctioned forms of assessment", while external assessment includes "all kinds of testing and assessment that is done because of external obligations from a school district, or from the state or federal government" (Stephens, 2004)

In this study, we adopted a newer and broader concept about assessment with focus on teachers' classroom assessment, or "internal assessment". Largely following the description given by the National Council of Teachers of Mathematics (NCTM, 1995) in its *Assessment standards for school mathematics*, we view assessment as:

> the process of teachers' gathering information about a student's knowledge of, ability to use, and disposition toward mathematics, mathematics teaching and mathematics learning, and of their making inferences from the information gathered for or abut students' learning in mathematics.

In general, assessment is essentially human activity of information-gathering and inference-making. In the domain of classroom instruction, assessment is the process of teachers' gathering information about students learning, which include their achievement and behavior in both cognitive and affective domains, and hence making informed decisions for classroom instruction; and the ultimate purpose of assessment is to improve the quality of teaching and learning. In relation to this, we view assessment as an essential part of the process of teaching and learning, and it should be integrated into classroom instructional activity.

Our focus in this research is on teachers' classroom assessment or internal assessment, as we believe that the integration of new assessment methods into mathematics classroom should to a large extent precede the use of those methods in external assessments. That is, the reform in classroom assessment should correspondingly lead to the reform in external assessment, not the other way around. By this way, desirable reforms in the high-stake external assessment can be better accepted and implemented (Fan, 2005).

Concerning alternative assessment strategies or, as we called in this study, new assessment strategies, there has been no universally agreed definition, nor has there been a universal classification, except that people generally agree that alternative assessment is different from traditional paper-and-pencil tests.

Nevertheless, for the last decades, the term alternative assessment in mathematics has been generally used to include the following specific techniques in student assessment: performance assessment, authentic assessment, portfolio assessment, journal writing assessment, project assessment, oral presentation assessment, interview assessment, classroom observation assessment, student self-assessment, student-constructed assessment, among others.

Each of the above-mentioned terms implies how the information was gathered in the process of assessment. For example, classroom observation assessment, or simply called classroom observation, refers to the assessment practice in which the teacher gathers the information about students' learning through observing their learning behavior and performance in classrooms.

This project focuses on four alternative assessment strategies: project assessment, performance assessment, student self-assessment, and communication assessment.

Performance assessment, or performance-based assessment, refers to the assessment practice in which the information about students' learning is gathered through students' work on performance tasks. Performance tasks in this study mainly include authentic real-life problems and open-ended tasks.

Project assessment, or project-based assessment, refers to the assessment practice in which the teacher gathers the information about students learning through their work on project tasks.

Student self-assessment refers to the assessment practice in which the information about students' learning is gathered through their reflection, evaluation, and report to the teacher.

Communication assessment in this study refers to the assessment practice in which the information about students' learning is gathered through students' performance on communication tasks, including mainly both journal writing (writing communication) and oral presentation (oral communication) tasks.

A detailed explanation about the concept of the four assessment strategies mentioned above is provided in relevant chapters from Chapter 3 to Chapter 10. Examples of different assessment tasks are also given in these chapters.

The reason that we focused on these four relatively new strategies is that, as we believe, not only are they better defined in the community of mathematics educators and practitioners, but also they have more practical importance and relevance to Singapore's educational system. For example, project work has been emphasized in Singapore school education in recent years (e.g., see Ministry of Education, 2001), and performance tasks have been believed to be helpful in developing students' problem-solving abilities and higher-order thinking skills. Moreover, student self-assessment requires students to take more responsibilities and be more engaged in self-reflection, and finally journal writing and oral presentation are of vital importance to the development of students' communication skills, which are much valued in the knowledge-based society. In short, these new assessment strategies relate to the areas in which Singaporean students are widely believed to be relatively weak and need improvement, compared to their western counterparts as revealed in large-scale international comparisons (e.g., see Ginsburg, Leinwand, Anstrom, & Pollok, 2005).

Readers who are interested to know more about assessment and alternative assessment may refer to the *Annotated Bibliography on Alternative Assessment in Mathematic* (Fan, Quek, Ng, et al., 2006) for a recent comprehensive review of the literature about the concept of assessment and alternative assessment. The bibliography was produced by this research team as a product of this project.

2.2 Participants of the Study

This study is a classroom-based one with invention of the new assessment strategies in classroom instruction for about three school semesters.

Eight secondary schools and eight primary schools participated in the study, and each of the four new assessment strategies was tried out in two secondary and two primary schools, respectively.

Using a stratified random sampling method, we first randomly selected eight primary and eight secondary schools from the population of all regular schools (primary: 174; secondary: 167) in the Singapore educational system, as then listed in the official MOE website (http://www.moe.gov.edu). Non-regular schools such as special education schools, international schools and sports schools were excluded.

At the primary level, four schools, known as high-performing primary schools in this study, were selected from Primary School Cohort I, which consisted of all the 9 schools with the top-level performance in terms of their students' average results in the 2003 Primary School Leaving Examination (PSLE) results, as provided by the Ministry in the following website:
http://www.getforme.com/previous2003/previous221103_psle2003results.htm. According to MOE, the following criteria were used for the identification of schools:

(a) The percentage of students eligible for secondary schools must be at least or above the national average of 97.2%;
(b) The school must have at least 80% of its students who are eligible for Special and Express course; and
(c) The percentage of students eligible for N(A), Express and Special courses must not dip by more than 2% compared to the 2002 results.

The other four primary schools, called non-high-performing primary schools in the study, were selected from Primary School Cohort II, the remaining schools.

At the secondary school level, it is well known that four types of courses (or called streams) have been provided in Singapore educational system with the main purpose to cater for the needs of students with different learning abilities. They are, from high to low in terms of academic levels, Special, Express, Normal Academic (NA), and Normal Technology (NT) courses. In this study, four high-performing secondary schools were randomly selected from Secondary School Cohort I, consisting of the 50 top performing schools, based on 2002 GCE O-Level Examinations Results for Special/Express Courses, as identified and released by MOE, and the other four non-high-performing secondary schools were from Secondary School Cohort II, the remaining schools.

After the schools were initially selected with the help of random function in a graphics calculator, we sent an invitation letter (see Appendix 2.1) to those schools for participating in this study. All the schools contacted accepted our invitation, except for a few schools which did not accept it due to different reasons (e.g., one school was already involved in a few classroom-based research projects). In addition, one secondary school initially expressed its willingness to participate, but after the project started it eventually decided not to do so because the teachers concerned felt there would be some difficulty, which is not easy to be dealt with, in carrying out the intervention in the way our project team hoped. Those schools were replaced with new schools selected, following the same criteria and procedures.

After the eight primary and secondary eight participating schools were identified, they were each randomly assigned to one new assessment strategy (see more detailed information below). After that, they were requested to nominate one to two high-performing classes and one to two non-high-performing classes at the grade level of Primary 3 and Secondary 1, respectively, for the project intervention, as well as one to two corresponding comparison classes. The general criteria used were, for the Primary 3 classes, the performance of the pupils[1] in the Primary 2 final examinations, and for the Secondary 1, according to students' overall PSLE results and course types—Special, Express, and Normal.

The grade levels of Primary 3 and Secondary 1 were selected for the commencement of intervention in order to avoid the streaming that schools may have for their students, as we noted that in primary and secondary schools, students would usually be streamed at the end of Primary 4 and Secondary 2 respectively. We also did not choose upper primary and secondary levels in order to avoid the extra pressure that teachers and students would likely face for the PSLE and O-level examinations which could cause potential problems and difficulties for them to carry out the intervention for this study, as it was designed to last for three semesters, In addition, we believe that, as Secondary 1 students just started their secondary education, they may be more willing to work on new methods and tasks in their learning; for Primary 3 pupils, they should be relatively familiar with the traditional assessment modes after two years of schooling and ready to accept new learning strategies. In fact, some researchers have argued that it is about this age that primary pupils may take to thinking and learning to learn (Nisbet & Shucksmith, 1984).

Table 2.1 summarizes the total number of the schools, classes, teachers, and students selected for the project invention.

Table 2.1. Number of the schools, classes, teachers, and students for invention

	Schools	Classes	Students	Teachers
Primary	8	15 (13)	625 (470)	16
Secondary	8	16 (15)	638 (590)	15
Total	16	31 (28)	1263 (1060)	31

Note. All the figures were collected at the beginning of the intervention. The figures in the brackets refer to comparison classes

A more detailed breakdown of participants and schools is presented in Table 2.2. As we can see, there were altogether 55 mathematics teachers and 2323 students involved for invention and comparison purpose. Note that, in Table 2.2 and hereafter, NP and HP stand for non-high-performing and high-performing primary schools, and NS and HS stand for non-high-performing and high-performing secondary schools, respectively. In addition, HC stands for "high-performing classes", whereas NHC stands for "non-high-performing classes".

Table 2.2. Numbers of participating schools and students

	Primary school								Secondary school								
	Cohort I				Cohort II				Cohort I				Cohort II				
	HP1	HP2	HP3	HP4	NP1	NP2	NP3	NP4	HS1	HS2	HS3	HS4	NS1	NS2	NS3	NS4	Total
HC																	
Int.	40	42	76	82	--	40	41	--	38	40	40	40	40	38	40	41	638
Com.	40	40	39	41	--	40	39	--	40	38	40	39	40	39	40	41	556
NHC																	
Int.	40	36	--	--	79	40	31	78	40	40	42	41	40	39	40	39	625
Com.	40	36	--	--	76	40	--	39	39	41	--	39	39	39	39	37	504
Total	160	154	115	123	155	160	111	117	157	159	122	159	159	155	159	158	2323

Note. Int.: "intervention classes" and Com.: "comparison classes". Due to various reasons, some schools did not have enough classes to be chosen as intervention classes or comparison classes, e.g., one high-performing secondary school (HS3) had only one non-high-performing class at Secondary One, so after it was chosen as intervention class, there was no comparison class. The figures in the table were collected at the beginning of the intervention.

Table 2.3 shows the information about the match between new assessment strategies and the participating schools. A research sub-team of two to four members were formed for the implementation of each of the four new assessment strategies, or simply called sub-study or sub-project, in both primary and secondary schools.

Table 2.3. Assignment of schools to assessment strategies

	Primary schools	Secondary schools
Performance Assessment	HP1, NP1	HS1, NS1
Communication Assessment	HP2, NP2	HS2, NS2
Project Assessment	HP3, NP3	HS3, NS3
Student Self-assessment	HP4, NP4	HS4, NS4

2.3 Instruments

Four main instruments were designed in the MAP project: pre- and post- questionnaire surveys, pre- and post- "new strategy task" tests, "new strategy" intervention tasks (and classroom observations), and interviews with students and teachers.

2.3.1 Questionnaires

Four sets of questionnaires were designed for the study. Two of them (pre-intervention and post-intervention) were for primary students and the other two (pre-intervention and post-intervention) for secondary students. The pre-intervention questionnaires were administered before the intervention in early 2004, whereas the post-intervention questionnaires were administered after the intervention in mid 2005.

The main purpose of the questionnaires was to measure the participating students' attitude towards mathematics and mathematics learning before and after intervention, and their experience with the new assessment strategies before and during the intervention. To a large extent, the primary and secondary versions of the questionnaires were designed to be parallel in terms of both content and format.

Both the pre- and post-intervention surveys, or simply pre-survey and post-survey, comprise two parts. The first part, identical in both surveys, focused on students' perceptions about the subject of mathematics and the value of mathematics learning. Four broad aspects were covered:

- General view towards mathematics and mathematics learning,
- Anxiety level in mathematics learning,
- Perceptions of their own performance in mathematics, and
- Beliefs about the usefulness of mathematics.

The only difference between the primary and secondary versions was in the number of questions (Primary: 14; Secondary: 22) and the phrasing of some questions in the consideration of students' comprehension abilities. For both school levels, the questionnaires used the Likert-type scale, and a nine-point scale ranging from "disagree totally" to "agree totally" was used for this part in the finalized versions. It was the project team's belief, that such a design is easier for the researchers to detect possible changes in students' attitudes before and after intervention.

The second part of pre-intervention questionnaires in both the primary and secondary versions was the same, containing a total of six questions with a six-point scale on frequency. These questions aimed to measure students' experience with the new assessment strategies in their previous mathematics learning. More specifically, one of the questions, including four sub-questions, asked the students about their experience with writing and speaking mathematics, two questions were about students' experience with self-reflection and self-construction, one question was about their experience with real-life problems, and the remaining two questions were on problems having various ways to solve or multiple correct answers. The corresponding part in the post-intervention one, still with a six-point Likert Scale, focused on students' feeling about using new assessment tasks in math learning.

A total of 79 Primary Three students selected from the population but not included in the sample took a pilot version of the survey. The pilot survey showed that the questionnaire could be completed within 15 minutes with and without teachers' assistance (with reading the items). Overall, students had no difficulty in answering the questionnaires with both scales. For some phrasing which some students had problems in understanding, a modification was made in the final version of the survey. A reliability test was done for the first part of questionnaire. To enhance the reliable level, four questions in the pilot version were finally removed.

Similarly, a pilot survey was conducted at the secondary level with 65 students from two schools from the population but not the sample participating in the pilot. Two questions were finally removed from the pilot version in order to enhance the reliable level and the phrasing of some questions was modified so as to make the questions more understandable to the students.

A copy of each pre-invention questionnaire mentioned can be found in Appendices 2.2 and 2.3. Note for post-invention questionnaire, its first part is the same as the pre-invention one. The second part of the post questionnaire for each sub-project is different, aiming to find out the students' opinions and experience on the particular assessment strategy after the intervention. More information about the second part for each assessment strategy can be found in the corresponding chapter in this report.

2.3.2 "New strategy task" tests

Similar to the questionnaire surveys, two sets of parallel "new strategy task" tests, a pre-test and a post-test, were designed and administered.

The purpose of the pre-test was for the researchers to obtain baseline measures on how well the students would perform on the new types of tasks so that researchers could understand better about the students' entry level as well as benefit from them in the designing subsequent interventions. The post-test was set to help the researchers to detect possible changes of students' ability in solving the new types of tasks after about three school terms with or without being exposed to the intervention tasks in mathematics learning.

Each research sub-team designed their own tests which were then reviewed and critiqued by the whole research team. The post-test tasks were designed to be parallel to the pre-test tasks.

A pilot-test was conducted for the pre-test with four classes of Primary Three students from one school. The tests lasted about 20 minutes, followed by a 10-minute question-and-answer session. The pilot test revealed the test questions to be new to the majority of students. Based on students' feedback from the pilot, some questions were rephrased to fit students' language abilities. Similarly, the "new strategy task" tests at the secondary level were piloted with five classes of Secondary One students from two schools. The results suggested that the secondary students were also not familiar with these new types of "assessment" questions. Some students also indicated that they had difficulty in understanding certain questions and corresponding modifications were made based on the feedback in the final version of the tests.

2.3.3 "New strategy" intervention tasks

As an intervention-based study, the implementation of intervention was essential to the project. In order to integrate new assessment strategies into the teachers' daily teaching practices, each research sub-team worked the new assessment tasks into the Schemes of Work (SOW) of the school under their charge[2].

Since in general the classes receiving intervention in each participating school are at different performance levels, different versions of the assessment tasks were often needed to match the students' ability and needs. Therefore, the researchers also worked with the participating teachers to confirm the appropriateness of the assessment tasks. The tasks were often presented in the format of worksheets.

To prepare and facilitate the implementation of the intervention strategies/tasks, two workshops, one at the beginning (February 2004) and the other in the middle (November 2004) of intervention were organized by the project team, for the participating teachers and their head of department. Some Vice-principals and Principals also attended the workshops. The first workshop consisted of an overall introduction about assessment and alternative assessment by the principal investigator, and in-depth sub-team meetings on the details of each of the individual assessment strategies, while the second one consisted of an overall review by the principal investigator about the progress made, presentation by some participating teachers on their experience (including challenges and difficulties) in implementing the new assessment strategies in the earlier stage of intervention, and sub-group discussion about the plan for the next stage of intervention.

Classroom observations were made by the researchers in each sub-team on most interventions with the main purpose to gather first-hand data on how the intervention tasks were carried out in the classrooms (many were audio- or video- taped). The classroom observations were also useful for the researchers to improve the design of new intervention tasks for the next stage. Field-notes and other anecdotal data were also collected during observations.

2.3.4 Interviews with teachers and students

To understand better participating teachers and students' experience, views, and suggestions about the use of the new assessment strategies, interviews with both the teachers and students of the intervention classes were conducted in May 2005, after the intervention was completed.

In general, about six students, two high-performing, two average-performing, and two low-performing, from each intervention class were randomly selected with the help of their teachers for the interview purpose. They were in most cases interviewed in two groups of threes, in two different sessions of about 20 to 40 minutes. The researchers felt that this arrangement would provide a more supportive and encouraging environment for them to respond to the interview questions. Almost all the participating teachers received an individual interview for about 30 to 60 minutes.

The interview questions were designed into two main groups. The first group of questions focused on the interviewees' own experience and understanding about the new assessment strategies implemented in their teaching and learning practice, for example, how they worked on the new assessment tasks and what difficulties they encountered. The second group focused on their personal views or suggestions about the use of the new assessment strategies based on their experiences, for example, how they feel about their new experiences and what they think can be further improved. The interview questions for both teachers and students are quite similar, except those for teachers are more from a teaching perspective and those for students are more from a learning perspective.

A copy of the general pre-structured interview questions and guidelines for teachers and students can be found in Appendices 2.4 and 2.5, respectively. For each sub-study, the sub-team members need to adjust the questions accordingly according to the new assessment strategy tasks (new tasks) and interventions they were involved.

2.3.5 School-based exams

Given the importance of school-based standard exams in local education contexts, the researchers were also interested to see how the new assessment strategies would impact students' learning as measured in these standard exams. Therefore, students' regular school-based classroom exams were also used as instruments/benchmarks to measure students' achievement in the cognitive domain during the period of intervention.

In all the sub-teams for the primary schools, the data or scores of students' regular school-based exams were collected from the exams that they took at the end of P2 (Oct/Nov 2003), before the intervention, in mid-P3 (May 2004), at the end of P3 (Oct/Nov 2004), and in mid-P4 (May 2004), which was at the end of the intervention.

In all the sub-teams for the secondary schools, the data or scores of students' regular school-based exams were collected from the exams that they took at the end of P6 (i.e., PSLE), before the intervention, in mid-Sec. 1 (May 2004), at the end of Sec. 1 (Oct/Nov 2004), and in mid-Sec. 2 (May 2004), which was again at the end of the intervention.

2.4 Data Collection and Analysis

As mentioned earlier, the pre-intervention questionnaire surveys were conducted in the 16 sample schools in early 2004, most in February and March, with a response rate being 97.6% for the primary schools and 97.6% for the secondary schools. The post-intervention questionnaire surveys were also conducted in May 2005 with an average response rate of more than 90% (see detailed figures in Chapters 3 to 8 respectively). We think that the high response rates were achieved mainly because of the schools' support for this research project.

The "new strategy task" pre-tests were conducted in the schools in March-April 2004, and the post-test were conducted in the schools in May 2005.

As mentioned earlier, classroom observations were made by team members throughout the intervention period, and interviews with students and teachers were conducted at the end of the intervention period. Many classroom observations, especially for the lessons involving the use of intervention tasks were audio- or video-taped, and all interviews were audio-taped.

In general, the data from the pre- and post- questionnaires were analyzed using quantitative methods. The descriptive analysis methods (e.g., frequency and percentage) were employed to describe students' overall perceptions about math as a subject and their learning of math before and after they participated in the study. When appropriate, Mann-Whitney U tests were used to examine the possible differences between the experimental classes and their corresponding comparison classes before and after the intervention period so as to enable researchers to detect the impact of using new assessment tasks on the experimental students' attitudes.

Students' work in the pre- and post- "new strategy task" tests was graded based on task-specific rubrics. When appropriate, the inter-reliability was calculated by the Intraclass Correlation Coefficient (ICC) on absolute agreement. Similar to the analysis for the questionnaire data, the rubric-based grades from the "new strategy task" tests were analyzed by descriptive statistics to investigate students' overall performance at class levels before and after intervention period. Mann-Whitney U test was employed to identify possible differences between the experimental and comparison classes in the two tests and school-based standard exams. Wilcoxon Signed-Ranks test was used to detect the change in students' grades from the pre- to post-tests. Some other more advanced statistical tools, such as significant tests and ANOVA, were also used to analyze the quantitative data collected.

All the interview data collected were transcribed by researchers themselves or in some cases by professional service providers, and examined by the researchers. Qualitative methods were employed in analysis. They were also used to triangulate what have been revealed in the quantitative data, so as to strength the findings of the study.

Chapter 3 Results and Findings (I): Performance Tasks (Primary)[3]

3.1 Introduction

In this chapter[4], we focus on the performance assessment at the primary school level. As mentioned in Chapter 2, performance assessment, or performance-based assessment, refers to the assessment practice in which the information about students' learning is gathered through students' work on performance tasks. In particular, the performance tasks used in this sub-project were all open-ended and, in many cases, authentic.

This component of the study investigated how the use of performance tasks impacted both attitudes and performance on various components of mathematics, including approach to a problem, solution to the problem and the way a solution is presented. By integrating performance tasks in mathematics teaching and learning, we hope that it is not only valuable for the students to develop concepts and problem solving skills, but also helpful for them to see the value of mathematics in the real world

3.2 Definition and Research Questions

3.2.1 Definition of terms

For the purpose of this project, performance tasks are operationally defined as open-ended tasks, which can often be “solved” in class within 2 periods (1 hour).

In addition to being open-ended, many tasks were also authentic. By authentic, we mean the problem contexts using either real-world data or the tasks involved real-work situations. It does not mean realistic in the sense that the last would necessarily be carried out by the student in their everyday life.

3.2.2 Research questions

As discussed in the introduction, the project focused on the impact of performance tasks over a period of about 18 months. The specific questions were:

(a) What is the impact of performance tasks on the attitude of students towards mathematics and performance tasks?
(b) Does project work impact the approach, presentation and ability to solve open-ended performance tasks?
(c) How does solving performance tasks impact students’ performance on traditional semestral mathematics tests?

As indicated earlier, it is our belief that assessment is a key component of mathematics education and should, therefore, be and integral part of a student’s mathematics education. Consequently the tasks are integrated into the classroom instruction. They are intended to be seen as part of the teaching and learning of mathematics; not as “add-ons” to the regular instruction.

3.3 Methods

3.3.1 Schools and classes

Two primary schools participated in this research on the use of performance tasks in mathematics. One is a high-performing school and the other, a typical neighbourhood school. Four classes were selected from each school, two experimental and two comparison classes. The information on these schools can best be summarized in Table 3.1.

Table 3.1. Participating schools and classes in performance assessment (Primary)

Class	School A: High-performing[1,2]	School B: Neighbourhood[1,2,3]
Experimental 1	High ability class in the school (HA) **E1**	Regular class (HA) **E3**
Experimental 2	“Low in terms of school” but good class relative to school B (LA) **E2**	Regular class (LA) **E4**
Comparison 1	Comparison Class for HA **C1**	Comparison Class **C3**
Comparison 2	“Low in terms of school” but good class relative to school B **C2**	Comparison Class **C4**

Note. [1]Class sizes varied from year to year, since some students left at the end of the first year and others joined. However, all classes had about 40 students, although the number for analysis was smaller (see Section 3.4 for more details). [2]All classes in School B are not of the level of School A. [3]There are no real differences between the two comparison classes.

3.3.2 Teachers

The participating teachers are all trained teachers. For School A there were no changes of teachers throughout the study. For School B the teacher of the low ability experimental class changed. Table 3.2 shows the information.

Table 3.2. Participating Teachers in Performance Assessment (Primary)

Initial teaching qualifications/Experience					
School A Highest qualification and experience			School B Highest qualification and experience		
E1	Dip Ed	2 years 3 months	E3	PGDE[1]	3 years 6 months
E2	MEd[1]	20 years	E4	Dip	15 years
C1	Cert Ed	23 years	C3	PGDE[1]	4 years
C2	FPDE[1]	15 years 9 months	C4	Dip Ed	1 year 6 months

Note. [1]These are postgraduate qualifications following undergraduate degrees.

As will be noted there is considerable variation in the background and experience of the teachers. It is worth noting that one of the experimental classes in each school was taught by a relatively new teacher.

3.3.3 Pre-tests and post-tests

Pre-tests
In this sub-project, pre-tests consisted both of a questionnaire survey of attitudes towards mathematics, as mentioned in Chapter 2 (see Appendix 2.1) as well as a test consisting of three performance tasks. They were conducted in all the 8 classes (4 experimental and 4 comparison) in the two primary schools.

The objective of the pre-tests was to provide baseline information on both attitudes and on students' ability to approach, present solutions and solve open-ended problems. The three test items can be found in Appendices 3.1. The rubric used to mark the performance task is in Appendix 3.2.

Post-tests
The post- tests were conducted 18 months later at the end of the experimental period. Both the questionnaire survey and performance task tests were designed to be parallel to the pre-tests. The only additional component was that the post-test attitude questionnaire contained an addition section asking for students' opinions on mathematics performance tests. The test items can be found in Appendix 3.3. The rubric used to mark the performance task is in Appendix 3.4.

3.3.4 Interviews

Student interviews
At the end of intervention, 6 students from each experimental class were interviewed using a set of structured interview questions as the basis. Parallel structured interview questions were used for all the components of the MAP project, with the expectation that each researcher would adapt the questions as appropriate.

The students were interviewed in pairs since it was felt that this would be a more supportive environment for the students. The goal was to provide additional information on their perceptions of performance tasks as well as triangulate the attitude questionnaire data.

Teacher interviews
At the end of the intervention the teachers involved with the experimental class were also interviewed using a set of structured interview questions as the basis. Again they were

interviewed in pairs in this sub-project. The goal was to provide information on their perceptions of performance tasks that would help in terms of any recommendations that might be useful regarding the use of performance tasks within the Singapore system.

3.3.5 Instructional approach

As mentioned earlier, the objectives of the intervention were for the teachers in the experimental classes to integrate the performance tasks into their regular teaching so that the students would see performance tasks as an integral component of instruction, not as an "add on to regular classroom instruction and assessment".

3.4 Results

3.4.1 Data set

As with any school-based research conducted over an extensive period there were changes in the student sample. At the end of primary 3, some students left the school while others changed classes.

The data analysis was only undertaken on those students who were involved for the total 18 months as well as were present for both the pre- and post- testing. This resulted in data being analyzed for the following number of students (Table 3.3).

Table 3.3. No. of students for data analysis (Performance assessment: Primary)

School A		School B	
Class	No. of students	Class	No. of students
E1	30	E3	34
E2	33	E4	34
C2	35	C3	32
C2	35	C4	34

3.4.2 Performance task test results

The performance tasks used in the pre- and post-tests are in Appendices 3.1 and 3.3. The readers will note that the post-test items parallel those in the pre-test. The major difference is in the second item on the tests. While the geometric activity is basically the same (making squares and making rectangles) the pre-test include graphical illustrations while the post-test did not.

As with any test involving problem solving, one cannot "guarantee" that they are totally parallel. Mathematically they are pairwise parallel. It would be noted that for any possibility that an item in the pre-test might be easier/harder than the corresponding item in the post-test, while it would impact the pre/post test comparisons, it would not impact any judgments made with regard to performance between the experimental and comparison groups.

Performance on individual tasks will now be discussed.

Movie/Song [referred to as movie task for convenience]
In this task, students had to work out how many combinations were allowed under specific conditions. The data will be discussed in two stages. We will first give a discussion of the data at a "student" level, followed by a statistical analysis of the data.

Table 3.4 summarizes the change of performance between pre- and post-test for each student on approach, presentation and solution. A more detailed discussion was given after the table.

Table 3.4. Results on the movie task

Movie	Approach			Presentation			Solution		
	+	-	=	+	-	=	+	-	=
School A									
E1	16	9	5	6	7	17	9	8	13
E2	12	10	11	5	18	10	10	15	8
C1	27	2	6	8	2	25	25	1	9
C2	9	15	11	6	19	10	8	16	11
School B									
E3	11	16	7	5	18	11	7	12	15
E4	18	4	13	17	15	3	28	2	4
C3	5	14	13	0	24	8	2	24	6
C4	13	10	11	4	16	14	17	4	13

Note. "+" means that the student performed better on the post-test than the pre-test, "-" means that the student performed worse on the post-test than the pre-test, and "=" means that the student performed at the same level on both tests.

Approach data
The results show the same basic pattern for both schools. For all classes, the majority of students improved their approach or, at least, performed no worse. However, there is no support for the conclusion that more students improved in the experimental classes. In fact, the best improvement was in the comparison class for School A. However, it should be noted that even when the total performance was no worse, in some cases there were more students whose performance went down than improved.

Presentation data
The results show lightly different patterns for the two schools. In fact, in 4 classes, one experimental and one comparison in each school, the majority of students performed worse in the post-test. However, it should be noted that for Class C4, although the total performance was no worse there were more students whose performance went down than improved.

Solution data
For the solution, only one class, a comparison class in School B had the majority of students who performed worse in post-test. However, it should be noted that even when the total performance was no worse, in some cases there were more students whose performance went down than improved.

Overall comment on movie data
Based on this data alone there is no clear pattern that differentiates the experimental from comparison classes or schools. The only clear implication is that some students benefit while others do not appear to benefit from the experience.

Square/Rectangle [referred to as square task for convenience]
Similarly, Table 3.5 summarizes the change of performance between pre- and post-test for each student on approach, presentation and solution on the square task.

Table 3.5. Results on the square task

Square	Approach			Presentation			Solution		
	+	-	=	+	-	=	+	-	=
School A									
E1	12	3	15	12	6	12	4	12	14
E2	2	20	11	2	23	8	3	23	7
C1	9	7	19	10	14	11	9	15	11
C2	12	8	15	13	12	10	12	13	10
School B									
E3	4	13	17	3	22	9	4	12	18
E4	21	1	12	18	9	7	21	7	6
C3	0	24	8	0	27	5	0	30	2
C4	8	12	14	4	19	11	11	13	10

Note. "+" means that the student performed better on the post-test than the pre-test, "-" means that the student performed worse on the post-test than the pre-test, and "=" means that the student performed at the same level on both tests.

Approach data
The majority of students improved their approach or, at least, performed no worse for all classes. The exception was for E2 and C3, where the majority of students performed worse of the post-test than the pre-test. However, it should be noted that even when the total performance was no worse, in some cases there were more students whose performance went down than improved.

Presentation data
The results show lightly different results from the approach data. Students in E2 and C3 still did not perform as well as the post-test and this also applied to students in E3 and C4. However, it should be noted that for C1 although the total performance was no worse, there were more students whose performance went down than improved.

Solution data
Here, the results were parallel to the approach data. However, it should be noted that even when the total performance was no worse, in most cases there were more students whose performance went down than improved.

Overall comment on square data
Based on the data alone there is no clear pattern that differentiates the experimental from comparison classes or schools. The only clear implication is that some students benefit while others do not appear to benefit from the experience.

Numbers
Table 3.6 summarizes the change of performance between pre- and post-test for each student on approach, presentation and solution on the numbers task.

Table 3.6. Results on the numbers task

Numbers	Approach			Presentation			Solution		
	+	-	=	+	-	=	+	-	=
School A									
E1	4	8	18	5	3	22	11	14	5
E2	6	11	16	9	6	18	17	16	0
C1	4	5	26	8	1	26	17	12	6
C2	7	13	15	10	5	20	7	20	8
School B									
E3	3	9	22	6	6	22	12	16	6
E4	3	19	12	6	11	17	6	25	3
C3	1	23	8	2	18	12	2	26	4
C4	14	6	14	20	5	9	15	14	5

Note. "+" means that the student performed better on the post-test than the pre-test, "-" means that the student performed worse on the post-test than the pre-test, and "=" means that the student performed at the same level on both tests.

Approach data
The majority of students improved their approach or, at least, performed no worse for all classes. The exception was for E4 and C3, both from School B, where the majority of students performed worse of the post-test than the pre-test. However, it should be noted that even when the total performance was no worse, in most cases there were more students whose performance went down than improved.

Presentation data
Except for C3 the majority of students in all groups performed at least as well on the post-test as on the pre-test. However, it should be noted that only in one case E4 were the total performance was no worse, did more students perform worse than improved.

Solution data
Students in E4, C3 and C2 did not perform as well on the post-test and the pre-test. All as groups performed as well. However, it should be noted that even when the total performance was no worse, in some cases there were more students whose performance went down than improved.

Overall comment on numbers data
Based on this data alone there is no clear pattern that differentiates the experimental from comparison classes or schools. The only clear implication is that some students benefit while others do not appear to benefit from the experience.

Summary of Results on Performance Tasks
The following table presents a summary of the results of the performance of all groups on the three tasks discussed above.

Table 3.7. Change in total performance by class on tasks by approach, presentation and solution (A: Approach; P: Presentation; S: Solution)

	Movie			Square			Numbers		
	A	P	S	A	P	S	A	P	S
Improved	**E1** **E2** E3 E4 **C1** **C2** C3 C4	**E1** E4 **C1**	**E1** **E2** E3 E4 **C1** **C2** C3	**E1** E3 E4 **C1** **C2**	**E1** E4 **C1** **C2** C4	**E1** E4 **C1** **C2** C4	**E1** **E2** E3 **C1** **C2** C4	**E1** **E2** E3 E4 **C1** **C2** C4	**E1** **E2** E3 **C1** C4
Worse		**E2** E3 **C2** C3 C4	C4	**E2** C3 C4	**E2** E3 C3	**E2** E3 C3	E4 C3	C3	E4 **C2** C3

Note. “Improved” means that “better” plus “equal level of performance” > “worse”. Bolded entries are from School A.

As mentioned earlier, when the performance of the classes on approach, presentation and solution on each task was analyzed, there were no clear patterns that emerged. However, an analysis of the summary table does provide some overall patterns that are worth discussing. The points worth noting are, for convenience, stated in point form for each school prior to the discussion.

School A:
(a) E1 and C1 improved on all tasks and categories within the tasks;
(b) E2’s major problems seem to be with the square’s task where the class did worse on all 3 criteria;
(c) C2 improved in all but 2 of the 9 categories.

School B:
(a) E3 and E4 improved in all but 2 and 3 categories respectively;
(b) C3 and C4 did worse in most categories.

Discussion
The data show different patterns for the two schools. It appears that for School A, the high ability school, except for the square task for the lower ability experimental group they improved in terms of overall classroom performance in virtually all areas. This would imply that their experience over the 18 months had a positive impact. However, based on this alone, there is no evidence to support the conclusion that the experimental groups improvement was much better that that of the comparison groups.

For School B there is a different pattern. The level of improvement for the experimental groups was similar to the lower ability experimental group in School A, although the categories where they did not do as well were for focused on a single performance task. The major difference is in the comparison groups. It appears that both comparison groups did not do well in the post-tests as compared to any of the other groups. It would appear that for School B, the neighborhood school, the performance tasks did have an obviously positive impact on the performance relative to the comparison class.

To analyze this further, let us look at each class in terms on the number of students who improved or performed worse (Table 3.8); it eliminates those who stayed the same.

Table 3.8 can help us refine the conclusion above. For both schools, it would appear that even when there was an overall improvement in the classes performance (based on improvement in level plus no change) in many cases, there were often more students who performed at a lower level than an higher level of the post-test. The question is how this modifies any conclusion relating to the impact of performance tasks.

Table 3.8. Change in improvement versus worse only on tasks by approach, presentation and solution (A: Approach; P: Presentation; S: Solution)

	Movie			Square			Numbers		
	A	P	S	A	P	S	A	P	S
Improved*	**E1** **E2** E4 **C1** C4	E4 **C1**	**E1** E3 E4 **C1** C3	**E1** E4 **C1** **C2**	**E1** E4 **C2**	E4		**E1 E2** E3[1] **C1 C2** C4	**E2** **C1** C4
Worse	*E3* ***C2*** *C3*	***E1*** **E2** E3 **C2** C3 C4	***E2*** ***C2*** C4	**E2** *E3* C3 C4	**E2** E3 ***C1*** C3 *C4*	***E1*** **E2** E3 ***C1*** ***C2*** C3 *C4*	***E1*** ***E2*** *E3* E4 ***C1*** ***C2*** C3 *C4*	E4 C4	***E1*** *E3* E4 **C2** C3

Note. “Improved” means that more students “improved their performance” than did “worse”. Italicised and underlined entries indicate that while the overall class improved (due to the number of students whose level of performance did not change) there were more students who did “worse” than actually “improved”. Bolded entries are from School A. [1]Equal number “improved” and were “worse”.

The research could try and look at specific performance in one category (e.g. approach to the square problem); what was the numerical difference (e.g. how many did better or worse in the presentation on the square problem); when a student changes was it from a 1 to 3 or 2 to 3, etc. (e.g. how many students in the numbers problem improved 1, 2 or 3 levels or went down 1, 2 or 3 levels on solution). However, given the nature of the data this level of data mining does not seem valid. The only realistic statement would be that even if there is an overall improvement in a class’s performance, for some students their performance over the 18 months went down.

To summarize the results, it appears that the use of performance tasks in the case of this sub-project had a more positive impact in a neighbourhood school than a high performing school. However, even when the overall performance in a class improved, there were a substantial number of students whose performance declined.

3.4.3 Student survey results

One of the objectives of this project dealt with potential change in students’ attitude about mathematics (and mathematics learning). As explained earlier, in order to research this objective an attitude questionnaire pre-survey was administered to all classes to determine students’ initial attitude towards mathematics. At the end of the intervention period, an attitude questionnaire post-survey was administered. The first part of the post-survey was identical to the pre-survey for all classes, while the second part for the experimental groups asked questions directly related to the use of performance tasks.

Data from Common Component in Pre- and Post- Surveys
In discussing the data, we will focus on situations where there is a change of more than 10% in perspective in at least one of the three categories: disagree, neutral or agree.

Basically, the 10% criterion is used as the initial basis for discussion. It is realized that this is an "arbitrary" criteria but seems a reasonable starting point.

General attitude
The first six questions in Part I are about students' general attitude toward mathematics and mathematics learning. Table 3.9 reports the results for Question 1 in the surveys.

Table 3.9. Results from Question 1 (I enjoy doing mathematics)

School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	0 (0%)	1 (2.5%)	***39 (97.5%)***	40
E1	Post-test	2 (5.2%)	4 (10.3%)	***33 (84.7%)***	39
E2	Pre-test	1 (2.6%)	4 (10.3%)	34 (87.2%)	39
E2	Post-test	5 (13.1%)	3 (7.9%)	30 (79%)	38
C1	Pre-test	3 (7.5%)	5 (12.5%)	32 (80%)	40
C1	Post-test	2 (5.2%)	5 (12.8%)	32 (82%)	39
C2	Pre-test	***3 (7.7%)***	***0 (0%)***	***36 (92.3%)***	39
C2	Post-test	***7 (18.4%)***	***8 (21.1%)***	***23 (60.5%)***	38
School B					
Class	Pre- or Post- test	Disagree	Neutral	Agree	Total
E3	Pre-test	7 (18.5%)	0 (0%)	31 (81.5%)	38
E3	Post-test	4 (10.2%)	0 (0%)	35 (89.8%)	39
E4	Pre-test	3 (7.5%)	2 (5%)	35 (87.5%)	40
E4	Post-test	5 (13.2%)	3 (7.9%)	30 (79%)	38
C3	Pre-test	***2 (5.1%)***	2 (5.1%)	35 (89.8%	39
C3	Post-test	***6 (16.8%)***	1 (2.8%)	29 (80.7%)	36
C4	Pre-test	***3 (8.4%)***	0 (0%)	***33 (91.7%)***	36
C4	Post-test	***6 (17.2%)***	2 (5.7%)	***27 (77%)***	35

Note. Figures in shaded lines are pre-survey data. Bolded and italicised entries indicate a situation where there is a change of more than 10%. Please note that the above notations will be used in the remaining tables in this chapter.

The data shows that two comparison classes and one experimental class seem to enjoy mathematics less than before. The experimental class, E1, is the class where 9 students left for the gifted programme and it is likely that the students who joined, while good at mathematics they might not be as positive as those who left. The same conclusions can be made about the most of the classes (although the percentage changes are smaller). The main exception is E3 where the students are more positive.

Table 3.10 reports the results for Question 2 in the surveys.

Table 3.10. Results from Question 2 (I am not afraid of doing mathematics)

Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	2 (5.2%)	3 (7.7%)	34 (87.1%)	39
E2	Pre-test	3 (7.7%)	1 (2.6%)	35 (89.7)	39
E2	Post-test	4 (10.6%)	5 (13.2%)	29 (76.3%)	38
C1	Pre-test	2 (5%)	4 (10%)	34 (85%)	40
C1	Post-test	2 (5.2%)	3 (7.7%)	34 (87.1%)	39
C2	Pre-test	6 (15.3%)	0 (0%)	33 (84.7%)	39
C2	Post-test	6 (15.8%)	3 (7.9%)	29 (76.3%)	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	5 (13.5%)	4 (10.8%)	28 (75.6%)	37
E3	Post-test	9 (23.1)	1 (2.6%)	29 (74.3%)	39
E4	Pre-test	3 (7.5%)	2 (5%)	35 (87.5%)	40
E4	Post-test	6 (15.8%)	3 (7.9%)	29 (76.3)	38
C3	Pre-test	4 (10.2%)	6 (15.4%)	29 (74.3%)	39
C3	Post-test	6 (16.8%)	2 (5.6%)	28 (77.8%)	36
C4	Pre-test	9 (25.7%)	0 (0%)	26 (74.3%)	35
C4	Post-test	10 (27.8%)	1 (2.8%)	25 (69.4%)	36

Overall, there appear no major changes in the results from the pre- to the post- survey on this question.

Table 3.11 reports the results for Question 3 in the surveys.

Table 3.11. Results from Question 3 (I am sure I can learn mathematics well)

Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	0 (0%)	1 (2.5%)	39 (97.5%)	40
E1	Post-test	1 (2.6%)	2 (5.1%)	36 (92.2)	39
E2	Pre-test	1 (2.6%)	***2 (5.1%)***	***36 (92.2)***	39
E2	Post-test	3 (7.9%)	***5 (13.2%)***	***30 (78.9%)***	38
C1	Pre-test	1 (2.5%)	3 (7.5%)	36 (90%)	40
C1	Post-test	3 (7.7%)	3 (7.7%)	33 (84.6%)	39
C2	Pre-test	1 (2.6%)	***3 (7.7%)***	***35 (89.8%)***	39
C2	Post-test	1 (2.6%)	***7 (18.4%)***	***30 (78.9%)***	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	4 (10.5%)	2 (5.3%)	32 (84.3%)	38
E3	Post-test	2 (5.2%)	3 (7.7%)	34 (87.2%)	39
E4	Pre-test	***2 (5%)***	3 (7.5%)	***35 (87.5%)***	40
E4	Post-test	***5 (13.1%)***	5 (13.1%)	***28 (73.7%)***	38
C3	Pre-test	3 (7.7%)	3 (7.7%)	33 (84.6%)	39
C3	Post-test	5 (14%)	3 (8.3%)	28 (77.8%)	36
C4	Pre-test	3 (8.4%)	2 (5.6%)	31 (86.2%)	36
C4	Post-test	5 (14%)	3 (8.3%)	28 (77.8%)	36

The data suggest that two comparison classes and one experimental class seem to feel that that cannot deal with mathematics as well as before the intervention. The same conclusions can be made about the most of the classes (although the percentage changes are smaller). There is a slight exception for E3 but the percentage change is very small.

Table 3.12 summarizes the results for Question 4 in the surveys.

Table 3.12. Results from Question 4 (I can get good grades in mathematics)

School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	0 (0%)	0 (0)%)	***40 (100%)***	40
E1	Post-test	1 (2.6%)	5 (12.8%)	***33 (84.5)***	39
E2	Pre-test	1 (2.6%)	5 (12.8%)	***33 (84.5)***	39
E2	Post-test	6 (15.7%)	7 (18.4%)	***25 (65.8%)***	38
C1	Pre-test	***0 (0%)***	4 (10.%)	36 (90%)	40
C1	Post-test	***5 (13.1%)***	2 (5.3%)	31 (81.6%)	38
C2	Pre-test	***2 (5.2%)***	***1 (2.6%)***	***36 (92.3%)***	39
C2	Post-test	***6 (15.8%)***	***9 (23.7%)***	***23 (60.6%)***	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	10 (26.3%)	4 (10.5%)	***24 (63.2%)***	38
E3	Post-test	8 (20.6%)	2 (5.1%)	***29 (74.3%)***	39
E4	Pre-test	6 (15%)	5 (12.5%)	***29 (72.5%)***	40
E4	Post-test	13 (34.2%)	3 (7.9%)	***22 (57.9%)***	38
C3	Pre-test	3 (7.8%)	8 (21.1%)	27 (71.1%)	38
C3	Post-test	10 (27.8%)	3 (8.3%)	23 (63.9%)	36
C4	Pre-test	4 (11.2%)	3 (8.3%)	29 (80.5%)	36
C4	Post-test	7 (20.1%)	3 (8.6%)	25 (71.4%)	35

We can observe from the table that there are different results between the two schools. For School A there is a clear pattern that students do not feel they can do as well in mathematics as they did before the intervention. This applies to both comparison and experimental classes.

For School B, there is no overall pattern. There are changes in the two experimental classes. For E3, the students feel they can get better grades while for E4, the reverse is true.

Table 3.13 lists the results for Question 5 in the surveys.

Table 3.13. Results from Question 5 (I think mathematics is useful to me)

		School A			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	0 (0%)	1 (2.5%)	39 (97.5%)	40
E1	Post-test	0 (0%)	0 (0%)	39 (100%)	39
E2	Pre-test	1 (2.6%)	1 (2.6%)	37 (94.8%)	39
E2	Post-test	0 (0%)	5 (12.8%)	34 (87.2)	39
C1	Pre-test	2 (5%)	2 (5%)	36 (90%)	40
C1	Post-test	0 (0%)	4 (10.5%)	34 (89.5%)	38
C2	Pre-test	3 (7.7%)	1 (2.6%)	35 (89.7%)	39
C2	Post-test	0 (0%)	1 (2.6%)	37 (97.4%)	38
		School B			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	***10 (26.4%)***	2 (5.3%)	***26 (68.4%)***	38
E3	Post-test	***2 (5.1%)***	2 (5.1%)	***35 (89.8%)***	39
E4	Pre-test	2 (5%)	2 (5%)	36 (90%)	40
E4	Post-test	4 (10.8%)	1 (2.70%)	32 (86.5%)	37
C3	Pre-test	3 (7.7%)	2 (5.1%)	34 (87.1%)	39
C3	Post-test	4 (11.2%)	0 (0%)	32 (89%)	36
C4	Pre-test	3 (8.4%)	4 (11.1%)	29 (80.6%)	36
C4	Post-test	4 (11.2%)	1 (2.8%)	31 (86%)	36

Potentially this is one area where the researchers would hope that the use of performance tasks, or the intervention, would have a positive impact. Clearly, again in the E3 there has been a large improvement in the class perspective of the usefulness of mathematics. For the other classes there has not been any change. It is likely that this is due to the fact that prior to the intervention they thought mathematics was useful; there was very little room for improvement.

Table 3.14 presents the results for Question 6 in the surveys.

Table 3.14. Results from Question 6 (Mathematics is interesting to me)

		School A			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	0 (0%)	1 (2.5%)	39 (97.5%)	40
E1	Post-test	1 (2.6%)	1 (2.6%)	37 (94.8%)	39
E2	Pre-test	4 (10.3%)	3 (7.7%)	32 (82.2%)	39
E2	Post-test	4 (10.5%)	4 (10.5%)	30 (78.9)	38
C1	Pre-test	3 (7.5%)	4 (10%)	33 (82.5%)	40
C1	Post-test	1 (2.6%)	4 (10.3%)	34 (87.2%)	39
C2	Pre-test	2 (5.1%)	***2 (5.1%)***	***35 (89.0%)***	39
C2	Post-test	3 (7.9%)	***7 (18.4%)***	***28 (73.7%)***	38
		School B			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	4 (10.4%)	2 (5.3%)	32 (84.3%)	38
E3	Post-test	3 (7.7%)	1 (2.6%)	35 (89.6%)	39
E4	Pre-test	3 (7.5%)	2 (5%)	35 (87.5)%	40
E4	Post-test	5 (13.1%)	2 (5.3%)	31 (81.6%)	38
C3	Pre-test	3 (7.7%)	3 (7.7%)	33 (84.6%)	39
C3	Post-test	5 (13.9%)	4 (11.1%)	27 (75%)	36
C4	Pre-test	4 (11.1%)	4 (11.1%)	28 (77.9%)	36
C4	Post-test	4 (11.1%)	6 (16.7%)	26 (72.3%)	36

The data in Table 3.14 show that there is little change in the interest level of mathematics for any of the classes. The exception is for C2, one of the comparison classes in School A where the percentage of students interested in mathematics has declined. For the other classes the changes are very small and they are varied in directionality.

Table 3.15 is about the results for Question 7 in the surveys.

Table 3.15. Results from Question 7 (Mathematics is hard for me)

		School A			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	***39 (97.5%)***	1 (2.5%)	***0 (0%)***	40
E1	Post-test	***28 (71.7%)***	4 (10.3%)	***7 (18%)***	39
E2	Pre-test	***27 (67.3%)***	***2 (5.1%)***	10 (25.7 %)	39
E2	Post-test	***18 (47.4%)***	***2 (18.4%)***	13 (34.2%)	39
C1	Pre-test	***31 (79.5%)***	4 (10.3%)	***4 (10.3%)***	39
C1	Post-test	***23 (58.9%)***	3 (7.7%)	***13 (33.3%)***	39
C2	Pre-test	27 (69.3%)	5 (12.8%)	7 (17.9%)	39
C2	Post-test	23 (60.7%)	6 (15.8%)	9 (23.7%)	38
		School B			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	21 (55.3%)	2 (5.3%)	15 (39.5%)	38
E3	Post-test	23 (58.9%)	3 (7.7%)	13 (33.3%)	39
E4	Pre-test	21 (52.2%)	2 (5%)	17 (42.5%)	40
E4	Post-test	21 (55.3%)	4 (10.5%)	13 (34.2%)	38
C3	Pre-test	13 (33.3%)	9 (23.1%)	17 (43.7%)	39
C3	Post-test	14 (38.9%)	4 (11.1%)	18 (50%)	36
C4	Pre-test	19 (52.7%)	2 (5.6%)	***15 (41.7%)***	36
C4	Post-test	16 (44.5%)	1 (2.8%)	***19 (52.8%)***	36

In School A, the high-performing school, it appears that the percentage of students who disagree with the statement that they find mathematics hard has decreased (this is also true for C2 where the percentage change is just under 10%). It seems reasonable to conclude that by the middle of Primary 4 mathematics in harder.

There is a different pattern in School B. Only in C4 does it appear that mathematics is harder.

It may well be that the expectations for all classes in the high achieving school could be the basis for this being for all the classes. It does not appear to be impacted by the experimental treatment. It should be noted that for E1 virtually all students started off positive, therefore the results could not improve further. Also, 9 students left this class to join the gifted education programme so the new students are unlikely to be academically as good.

Table 3.16 is about the results for Question 8 in the surveys.

Table 3.16. Results from Question 8 (Knowing mathematics will help me get a good job next time)

School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	2 (5%)	5 (12.5%)	***33 (82.5%)***	40
E1	Post-test	0 (0%)	2 (5.1%)	***37 (94.9%)***	39
E2	Pre-test	8 (20.6%)	3 (7.7%)	28 (71.9%)	39
E2	Post-test	4 (10.4%)	6 (15.8%)	28 (73.8%)	38
C1	Pre-test	3 (7.7%)	***5 (12.8%)***	***31 (79.5%)***	39
C1	Post-test	2 (5.2%)	***13 (33.3%)***	***24 (61.5%)***	39
C2	Pre-test	3 (7.7%)	9 (23.1%)	27 (69.2%)	39
C2	Post-test	2 (5.3%)	11 (28.9%)	25 (65.8%)	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	4 (10.5%)	4 (10.5%)	30 (79%)	38
E3	Post-test	1 (2.6%)	5 (12.8%)	33 (84.7%)	39
E4	Pre-test	4 (10%)	1 (2.5%)	35 (87.5%)	40
E4	Post-test	1 (2.7%)	2 (5.4%)	34 (91.9%)	37
C3	Pre-test	3 (7.7%)	3 (7.7%)	33 (84.6%)	39
C3	Post-test	4 (11.1%)	3 (8.3%)	29 (80.6%)	36
C4	Pre-test	4 (11.2%)	5 (13.9%)	27 (74.9%)	36
C4	Post-test	7 (19.5%)	4 (11.1%)	25 (69.5%)	36

The table shows that all experimental classes become more positive, but all comparison classes become more negative about the value of mathematics in terms of getting a job, in terms of the percentage of students in each class answering “yes”. In particular, the change of percentage is more than 10% in E1, and also more than 10% but in negative direction in C1.

Table 3.17 reports the results of Question 9 in the surveys.

Table 3.17. Results from Question 9 (I don’t feel good about mathematics)

School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	***39 (97.5%)***	***1 (2.5%)***	***0 (0%)***	40
E1	Post-test	***28 (71.7%)***	***6 (15.4%)***	***5 (12.8%)***	39
E2	Pre-test	***30 (79%)***	2 (5.3%)	6 (15.9%)	38
E2	Post-test	***25 (65.8%)***	4 (10.5%)	9 (23.7%)	38
C1	Pre-test	32 (80%)	3 (7.5%)	5 (12.5%)	40
C1	Post-test	33 (84.6%)	2 (5.1%)	4 (10.2%)	39
C2	Pre-test	***33 (84.6%)***	***3 (7.7%)***	3 (7.7%)	39
C2	Post-test	***25 (65.8%)***	***9 (23.7%)***	4 (10.6%)	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	***21 (56.7%)***	3 (8.1%)	***13 (35.1%)***	37
E3	Post-test	***29 (74.4%)***	3 (7.7%)	***7 (18%)***	39
E4	Pre-test	28 (70%)	2 (5%)	10 (25%)	40
E4	Post-test	28 (73.6%)	4 (10.5%)	6 (15.8%)	38
C3	Pre-test	***18 (46.1%)***	***12 (30.8%)***	***9 (23.1%)***	39
C3	Post-test	***22 (61.2%)***	***2 (5.6%)***	***12 (33.3%)***	36
C4	Pre-test	***24 (66.6%)***	1 (2.8%)	11 (30.7%)	36
C4	Post-test	***20 (55.6%)***	4 (11.1%)	12 (33.3%)	36

Except for C1 and E4, the data presented in Table 3.17 show that for many classes their feel worse about mathematics after 18 months than they did prior to the start of the intervention. It seems that classes are more negative and this seems to be relatively independent on school and comparison/experimental classes.

Table 3.18 reports the results for Question 10 in the surveys.

Table 3.18. Results from Question 10 (It makes me nervous to do mathematics)

		School A			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	***38 (95%)***	1 (2.5%)	1 (2.5%)	40
E1	Post-test	***33 (84.6%)***	2 (5.1%)	4 (10.3%)	39
E2	Pre-test	28 (71.7%)	3 (7.7%)	8 (20.5%)	39
E2	Post-test	24 (63.1%)	4 (10.5%)	10 (26.3%)	38
C1	Pre-test	31 (77.5%)	1 (2.5%)	8 (20%)	40
C1	Post-test	32 (82.1%)	3 (7.7%)	4 (10.3%)	39
C2	Pre-test	28 (71.7%)	3 (7.7%)	8 (20.5%)	39
C2	Post-test	24 (63.1%)	7 (18.4%)	7 (18.4%)	38
		School B			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	***25 (67.5%)***	2 (5.4%)	***10 (27%)***	37
E3	Post-test	***21 (53.8%)***	3 (7.7%)	***15 (38.5%)***	39
E4	Pre-test	25 (62.5%)	6 (15%)	9 (22.5%)	40
E4	Post-test	20 (52.6%)	7 (18.4%)	11 (29%)	38
C3	Pre-test	***17 (44.8%)***	***10 (26.3%)***	***11 (28.9%)***	38
C3	Post-test	***20 (55.5%)***	***2 (5.6%)***	***14 (38.9%)***	36
C4	Pre-test	16 (44.5%)	6 (16.7%)	14 (38.9%)	36
C4	Post-test	17 (47.2%)	3 (8.3%)	16 (44.4%)	36

The table shows that For E1 and E3 the classes are slightly more nervous about mathematics. For C3 the trend is the same except more students are now neutral. Table 3.19 is about the results for Question 8 in the surveys.

Table 3.19. Results from Question 11 (I am not good at mathematics)

		School A			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	38 (95%)	***1 (2.5%)***	1 (2.5%)	40
E1	Post-test	29 (74.3%)	***6 (15.4%)***	4 (10.3%)	39
E2	Pre-test	***27 (69.2%)***	5 (12.8%)	***7 (18%)***	39
E2	Post-test	***19 (50%)***	4 (10.5%)	***15 (39.5%)***	38
C1	Pre-test	32 (80%)	5 (12.5%)	***3 (7.5%)***	40
C1	Post-test	29 (74.3%)	2 (5.1%)	***8 (20.5%)***	39
C2	Pre-test	***26 (68.4%)***	4 (10.5%)	***8 (21.1%)***	38
C2	Post-test	***19 (50%)***	4 (10.5%)	***15 (39.5%)***	38
		School B			
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	20 (52.7%)	2 (5.3%)	16 (42.1%)	38
E3	Post-test	21 (53.8%)	3 (7.7%)	15 (38.5%)	39
E4	Pre-test	***31 (77.5%)***	3 (7.5%)	***6 (15%)***	40
E4	Post-test	***22 (57.9)***	3 (7.9%)	***13 (34.2%)***	38
C3	Pre-test	16 (41.1)	7 (17.9%)	16 (41%)	39
C3	Post-test	14 (38.9%)	4 (11.1%)	18 (50%)	36
C4	Pre-test	***21 (58.4%)***	***7 (19.4%)***	***8 (22.3%)***	36
C4	Post-test	***14 (38.9%)***	***2 (5.6%)***	***20 (55.6)***	36

According to the results presented in the table, perceptions about being good at mathematics changed in all the classes in School A and in two of the classes in School B. Overall, the students do not believe they are as good at mathematics as there were 18 months ago, which is highly consistent with findings from other researches observing the change of students attitudes in mathematics from lower grades to higher grades as mathematics becomes harder, and hence quite understandable.

Table 3.20 shows the results for Question 12 in the surveys.

Table 3.20. Results from Question 12 (I dislike mathematics)

School A					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E1	Pre-test	***39 (97.5%)***	0 (0%)	1 (2.5%)	40
E1	Post-test	***33 (84.6%)***	2 (5.1%)	4 (10.3%)	39
E2	Pre-test	***33 (84.7%)***	1 (2.6%)	5 (12.9%)	39
E2	Post-test	***28 (73.7%)***	3 (7.9%)	7 (18.4%)	38
C1	Pre-test	35 (87.5%)	2 (5%)	3 (7.5%)	40
C1	Post-test	31 (81.5%)	4 (10.5%)	3 (7.8%)	38
C2	Pre-test	34 (87.2%)	4 (10.3%)	1 (2.6%)	39
C2	Post-test	34 (87.2%)	4 (10.3%)	1 (2.6%)	39
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	***24 (63.2%)***	3 (7.9%)	***11 (29%)***	38
E3	Post-test	***33 (84.6%)***	3 (7.7%)	***3 (7.7%)***	39
E4	Pre-test	26 (66.7%)	4 (10.3%)	9 (23.1%)	39
E4	Post-test	25 (71.5%)	2 (5.7%)	8 (22.8%)	35
C3	Pre-test	***24 (61.6%)***	***10 (25.6%)***	5 (12.8%)	39
C3	Post-test	***31(81.6%)***	***1 (2.6%)***	6 (15.7)	38
C4	Pre-test	24 (66.7%)	3 (8.3%)	9 (25.1%)	36
C4	Post-test	25 (69.5%)	4 (11.1%)	7 (19.5%)	36

The results show that both experimental classes in School A, the high achieving school, have a more negative view of mathematics at the end of 18 months. However, for one experimental and one comparison class in School B, the neighborhood school, they now have a more positive view of mathematics.

Table 3.21 is about the results for Question 13 in the surveys.

Table 3.21. Results from Question 13 (I like to do difficult mathematics questions)

Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
School A					
E1	Pre-test	***1 (2.5%)***	0 (0%)	***39 (97.5%)***	40
E1	Post-test	***7 (18%)***	3 (7.7%)	***29 (74.4%)***	39
E2	Pre-test	8 (20.5%)	5 (12.8%)	***26 (66.6%)***	39
E2	Post-test	11 (28.9%)	8 (21.1%)	***19 (50%)***	38
C1	Pre-test	7 (18%)	***3 (7.7%)***	29 (74.3%)	39
C1	Post-test	5 (12.9%)	***8 (20.5%)***	26 (66.6%)	39
C2	Pre-test	***7 (18%)***	5 (12.8%)	***27 (69.2%)***	39
C2	Post-test	***13 (34.2%)***	6 (15.8%)	***19 (50%)***	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	14 (36.8%)	2 (5.3%)	22 (57.8%)	38
E3	Post-test	13 (33.3%)	5 (12.8%)	21 (53.8%)	39
E4	Pre-test	***11 (27.5%)***	3 (7.5%)	***26 (65%)***	40
E4	Post-test	***19 (50%)***	4 (10.5%)	***15 (39.5%)***	38
C3	Pre-test	***8 (20.5%)***	6 (15.4%)	***25 (64.2%)***	39
C3	Post-test	***14 (38.8%)***	5 (13.9%)	***17 (47.2)%***	36
C4	Pre-test	***8 (22.3%)***	***6 (16.7%)***	***22 (61.1%)***	36
C4	Post-test	***19 (52.8)%***	***2 (5.6%)***	***15 (41.7%)***	36

The table shows that except for E3 there is a change in the perception of liking to do challenging questions. For the 7 classes where there is a change, 6 are more negative while E1 has the greatest percentage increase in those who are neutral. It appears that, in general, independent of school or comparison/experimental, students are less enamored of a challenge.

Table 3.22 is about the results for Question 14 in the surveys.

Table 3.22. Results from Question 14 (I like spending time on studying mathematics)

Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
School A					
E1	Pre-test	2 (5%)	***4 (10.%)***	***34 (85%)***	40
E1	Post-test	5 (12.8%)	***9 (23.1%)***	***25 (64.1%)***	39
E2	Pre-test	10 (25.6%)	***1 (2.6%)***	***28 (71.9%)***	39
E2	Post-test	11 (28.9%)	***8 (21.1%)***	***19 (50%)***	38
C1	Pre-test	8 (20.5%)	4 (10.3%)	27 (69.3%)	39
C1	Post-test	8 (20.5%)	4 (10.3%)	27 (69.2%)	39
C2	Pre-test	6 (15.4%)	***6 (15.4%)***	***27 (69.2%)***	39
C2	Post-test	5 (13.2%)	***18 (47.4%)***	***15 (39.4%)***	38
School B					
Class	Pre- or Post-test	Disagree	Neutral	Agree	Total
E3	Pre-test	5 (13.1%)	***2 (5.3%)***	***31 (81.6%)***	38
E3	Post-test	5 (12.8%)	***6 (15.4%)***	***28 (71.8%)***	39
E4	Pre-test	7 (17.5%)	***6 (15.%)***	27 (67.5%)	40
E4	Post-test	5 (13.2%)	***10 (26.3%)***	23 (60.6%)	38
C3	Pre-test	***5 (12.9%)***	4 (10.3%)	***30 (76.9%)***	39
C3	Post-test	***9 (25%)***	4 (11.1%)	***23 (64%)***	36
C4	Pre-test	6 (16.7%)	3 (8.3%)	27 (75%)	36
C4	Post-test	4 (11.1%)	6 (16.7%)	26 (72.2%)	36

The figures in the table show that in School A three of the four classes the percentage of students who like spending time studying mathematics has decreased. The same situation applies to E1 and C3 in School B.

Changes in Teaching: Comparison Groups
The second part of the questionnaire in the post-test was different for the experimental and comparison groups. For the comparison groups this part of the questionnaire was identical to the pre-test. The results will be explored to see if there is any change in approach for these groups. The 10% criteria will be used, although with more cells is it realized that this is a more difficult criteria to meet. We start with comparison groups.

Table 3.23 is about the results for Question 15A in the surveys.

Table 3.23. Results from Question 15A for comparison classes (My math teacher had asked me to write down the reasons for my math answer)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***8 (20.5%)***	1 (2.6%)	***0 (0%)***	1 (2.6%)	5 (12.8%)	***24 (61.5%)***	39
C1	Post-test	***13 (33.3%)***	5 (12.8%)	***7 (17.9%)***	1 (2.6%)	4 (10.3%)	***9 (23.1%)***	39
C2	Pre-test	18 (46.2%)	2 (5.1%)	***4 (10.3%)***	0 (0%)	1 (2.6%)	***14 (35.9%)***	39
C2	Post-test	17 (44.7%)	4 (10.5%)	***8 (21.1%)***	0 (0%)	1 (2.6%)	***8 (21.1%)***	38
				School B				
C3	Pre-test	***22 (59.5%)***	6 (16.2%)	3 (8.1%)	3 (8.1%)	1 (2.7%)	***2 (5.4%)***	37
C3	Post-test	***12 (34.3%)***	9 (25.7%)	3 (8.6%)	0 (0%)	2 (5.7%)	***9 (25.7%)***	35
C4	Pre-test	9 (25.0%)	9 (25.0%)	5 (13.9%)	1 (2.8%)	4 (11.1%)	8 (22.2%)	36
C4	Post-test	7 (20.6%)	9 (26.5%)	5 (14.7%)	2 (5.9%)	1 (2.9%)	10 (29.4%)	34

It appears from the table that for C1 students are asked to write down their reasons far less than at the beginning of the intervention period. The same is true for C2 but to a lesser extent. For C3 the trend is in the opposite direction.

Table 3.24 is about the results for Question 15B in the surveys.

Table 3.24. Results from Question 15B for comparison classes (My math teacher had asked me to explain mathematics to the whole class)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***9 (23.1%)***	4 (10.3%)	3 (7.7%)	0 (0%)	4 (10.3%)	***19 (48.7%)***	39
C1	Post-test	***21 (53.8%)***	7 (17.9%)	1 (2.6%)	1 (2.6%)	5 (12.8%)	***4 (10.3%)***	39
C2	Pre-test	9 (23.7%)	***9 (23.7%)***	2 (5.3%)	1 (2.6%)	3 (7.9%)	14 (36.8%)	38
C2	Post-test	9 (23.7%)	***5 (13.2%)***	5 (13.2%)	4 (10.5%)	4 (10.5%)	11 (28.9%)	38
				School B				
C3	Pre-test	***21 (58.3%)***	2 (5.6%)	2 (5.6%)	***0 (0%)***	6 (16.7%)	***5 (13.9%)***	36
C3	Post-test	***9 (25.7%)***	4 (11.4%)	3 (8.6%)	***4 (11.4%)***	3 (8.6%)	***12 (34.3%)***	35
C4	Pre-test	***9 (25.7%)***	5 (14.3%)	2 (5.7%)	1 (2.9%)	3 (8.6%)	***15 (42.9%)***	35
C4	Post-test	***13 (37.1%)***	5 (14.3%)	5 (14.3%)	1 (2.9%)	1 (2.9%)	***10 (28.6%)***	35

The results show that in both C3 and C4 teachers ask them less frequently to explain to whole class. For C1 the teacher asks them to explain more. There is no major change in C2.

Table 3.25 reports the results for Question 15C in the surveys.

The data in Table 3.25 show that in two of the four classes there was more opportunity for students to write down their feelings towards mathematics (C1 and C3) while in C4 the trend is in the other direction.

Table 3.25. Results from Question 15C for comparison classes (My math teacher had asked me to write down my feelings about mathematics)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***4 (10.3%)***	1 (2.6%)	***4 (10.3%)***	***1 (2.6%)***	***0 (0%)***	***29 (74.4%)***	39
C1	Post-test	***0 (0%)***	1 (2.6%)	***0 (0%)***	***5 (12.8%)***	***9 (23.1%)***	***24 (61.5%)***	39
C2	Pre-test	3 (7.9%)	3 (7.9%)	3 (7.9%)	2 (5.3%)	3 (7.9%)	24 (63.2%)	38
C2	Post-test	0 (0%)	1 (2.6%)	3 (7.9%)	5 (13.2%)	5 (13.2%)	24 (63.2%)	38
				School B				
C3	Pre-test	2 (5.4%)	1 (2.7%)	1 (2.7%)	2 (5.4%)	2 (5.4%)	***29 (78.4%)***	37
C3	Post-test	5 (14.3%)	2 (5.7%)	2 (5.7%)	2 (5.7%)	2 (5.7%)	***22 (62.9%)***	35
C4	Pre-test	3 (8.3%)	1 (2.8%)	2 (5.6%)	2 (5.6%)	13 (36.1%)	***15 (41.7%)***	36
C4	Post-test	4 (11.8%)	1 (2.9%)	2 (5.9%)	1 (2.9%)	4 (11.8%)	***22 (64.7%)***	34

Table 3.26 is about the results for Question 15D in the surveys.

The data in Table 3.26 show that in C1 there is a large increase in students being asked to explain ideas in writing while the reverse is true of C4.

Table 3.26. Results from Question 15D for comparison classes (My math teacher had asked me to explain math ideas in writing)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***5 (12.8%)***	3 (7.7%)	4 (10.3%)	2 (5.1%)	***2 (5.1%)***	***23 (59.0%)***	39
C1	Post-test	***8 (20.5%)***	4 (10.3%)	5 (12.8%)	1 (2.6%)	***15 (38.5%)***	***6 (15.4%)***	39
C2	Pre-test	***9 (23.7%)***	7 (18.4%)	***3 (7.9%)***	1 (2.6%)	2 (5.3%)	16 (42.1%)	38
C2	Post-test	***5 (13.2%)***	8 (21.1%)	***7 (18.4%)***	1 (2.6%)	4 (10.5%)	13 (34.2%)	38
				School B				
C3	Pre-test	***27 (73%)***	3 (8.1%)	2 (5.4%)	***0 (0%)***	3 (8.1%)	2 (5.4%)	37
C3	Post-test	***7 (20.6%)***	5 (14.7%)	4 (11.8%)	***4 (11.8%)***	4 (11.8%)	10 (9.4%)	34
C4	Pre-test	12 (33.3%)	6 (16.7%)	4 (11.1%)	4 (11.1%)	5 (13.9%)	***5 (13.9%)***	36
C4	Post-test	8 (24.2%)	5 (15.2%)	5 (15.2%)	1 (3.0%)	1 (3.0%)	***13 (39.4%)***	33

Table 3.27 presents the results for Question 16 in the surveys.

Table 3.27. Results from Question 16 for comparison classes (My math teacher encouraged me to solve math questions in different ways)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***12 (31.6%)***	6 (15.8%)	***9 (23.7%)***	***0 (0%)***	2 (5.3%)	***9 (23.7%)***	38
C1	Post-test	***8 (20.5%)***	9 (23.1%)	***13 (33.3%)***	***5 (12.8%)***	2 (5.1%)	***2 (5.1%)***	39
C2	Pre-test	***14 (35.9%)***	14 (35.9%)	***1 (2.6%)***	3 (7.7%)	3 (7.7%)	4 (10.3%)	39
C2	Post-test	***8 (21.1%)***	14 (36.8%)	***6 (15.8%)***	3 (7.9%)	0 (0%)	7 (18.4%)	38
				School B				
C3	Pre-test	***21 (53.8%)***	9 (23.1%)	***3 (7.7%)***	***0 (0%)***	4 (10.3%)	***2 (5.1%)***	39
C3	Post-test	***9 (26.5%)***	6 (17.6%)	***6 (17.6%)***	***4 (11.8%)***	3 (8.8%)	***6 (17.6%)***	34
C4	Pre-test	***10 (27.8%)***	8 (22.2%)	***8 (22.2%)***	4 (11.1%)	3 (8.3%)	***3 (8.3%)***	36
C4	Post-test	***13 (40.6%)***	9 (28.1%)	***2 (6.3%)***	1 (3.1%)	0 (0%)	***7 (21.9%)***	32

The data in Table 3.27 show that in three of the four classes there is a decrease in the opportunity to solve questions in different ways while in C1 the trend is in the other direction.

Table 3.28 presents the results for Question 17 in the surveys.

Table 3.28. Results from Question 17 for comparison classes (My math teacher asked me to make up math questions by myself)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***6 (16.2%)***	1 (2.7%)	2 (5.4%)	0 (0%)	5 (13.5%)	23 (62.2%)	37
C1	Post-test	***2 (5.1%)***	0 (0%)	3 (7.7%)	2 (5.1%)	8 (20.5%)	24 (61.5%)	39
C2	Pre-test	***8 (20.5%)***	4 (10.3%)	4 (10.3%)	2 (5.1%)	5 (12.8%)	***16 (41.0%)***	39
C2	Post-test	***1 (2.6%)***	2 (5.3%)	3 (7.9%)	1 (2.6%)	5 (13.2%)	***26 (68.4%)***	38
				School B				
C3	Pre-test	10 (26.3%)	8 (21.1%)	3 (7.9%)	2 (5.3%)	4 (10.5%)	***11 (28.9%)***	38
C3	Post-test	7 (20.0%)	4 (11.4%)	3 (8.6%)	3 (8.6%)	1 (2.9%)	***17 (48.6%)***	35
C4	Pre-test	7 (19.4%)	7 (19.4%)	***6 (16.7%)***	0 (0%)	***6 (16.7%)***	10 (27.8%)	36
C4	Post-test	2 (5.7%)	6 (17.1%)	***1 (2.9%)***	1 (2.9%)	***13 (37.1%)***	12 (34.3%)	35

The data show that in C2 and C3 there was substantially less opportunity to make up their own questions, while in other two classes, there is no major change.

Table 3.29 presents the results of Question 18 in the surveys.

Table 3.29. Results from Question 18 for comparison classes (How often did your math teacher ask you think about the reason for your solving math problems?)

Class	Pre- or Post-test	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total
				School A				
C1	Pre-test	***6 (15%)***	***6 (15%)***	7 (17.5%)	2 (5%)	***2 (5%)***	***17 (42.5%)***	40
C1	Post-test	***11 (28.2%)***	***10 (25.6%)***	3 (7.7%)	5 (12.8%)	***6 (15.4%)***	***4 (10.3%)***	39
C2	Pre-test	13 (33.3%)	9 (23.1%)	5 (12.8%)	1 (2.6%)	1 (2.6%)	10 (25.6%)	39
C2	Post-test	13 (34.2%)	6 (15.8%)	5 (13.2%)	2 (5.3%)	2 (5.3%)	10 (26.3%)	38
				School B				
C3	Pre-test	***16 (41.0%)***	3 (7.7%)	5 (12.8%)	5 (12.8%)	3 (7.7%)	***7 (17.9%)***	39
C3	Post-test	***9 (25%)***	4 (11.1%)	3 (8.3%)	3 (8.3%)	6 (16.7%)	***11 (30.6%)***	36
C4	Pre-test	9 (26.5%)	9 (26.5%)	6 (17.6%)	1 (2.9%)	6 (17.6%)	***3 (8.8%)***	34
C4	Post-test	9 (25.7%)	7 (20%)	5 (14.3%)	4 (11.4%)	3 (8.6%)	***7 (20%)***	35

The results reveal that in C1 there was an increase in the opportunity to that about reasons for solving mathematics problems while the reverse was true for C2 and C3.

Table 3.30 presents the results for Question 19 in the surveys.

The figures in the table suggest that in C3 and C4 there was a decrease on the teacher asking the students to have more than one correct answer. The change in other two classes is less substantial.

Table 3.30. Results from Question 19 for comparison classes (How many math questions did your teacher ask you to do have more than 1 correct answer?)

Class	Pre- or Post-test	Almost all	More than half	Half	Less than half	A Few	None	Total
				School A				
C1	Pre-test	16 (41%)	9 (23.1%)	5 (12.8%)	3 (7.7%)	3 (7.7%)	3 (7.7%)	39
C1	Post-test	15 (38.5%)	12 (30.8%)	5 (12.8%)	2 (5.1%)	5 (12.8%)	0 (0%)	39
C2	Pre-test	18 (47.4%)	8 (21.1%)	6 (15.8%)	0 (0%)	6 (15.8%)	0 (0%)	38
C2	Post-test	16 (42.1%)	11 (28.9%)	4 (10.5%)	3 (7.9%)	2 (5.3%)	2 (5.3%)	38
				School B				
C3	Pre-test	***12 (30.8%)***	9 (23.1%)	11 (28.2%)	***2 (5.1%)***	5 (12.8%)	0 (0%)	39
C3	Post-test	***4 (11.1%)***	8 (22.2%)	12 (33.3%)	***6 (16.7%)***	4 (11.1%)	2 (5.6%)	36
C4	Pre-test	10 (29.4%)	***14 (41.2%)***	***2 (5.9%)***	1 (2.9%)	4 (11.8%)	3 (8.8%)	34
C4	Post-test	10 (28.6%)	***6 (17.1%)***	***11 (31.4%)***	3 (8.6%)	4 (11.4%)	1 (2.9%)	35

Table 3.31 is about the results for Question 20 in the surveys.

The results in the table show that in all classes there was a decrease in the mathematics questions relating to the real world.

Table 3.31. Results from Question 20 for comparison classes (How many math questions did your teacher ask you to do have nothing to do real life?)

Class	Pre- or Post-test	Almost all	More than half	Half	Less than half	A Few	None	Total
				School A				
C1	Pre-test	10 (26.3%)	5 (13.2%)	9 (23.7%)	5 (13.2%)	4 (10.5%)	5 (13.2%)	38
C1	Post-test	4 (10.3%)	2 (5.1%)	6 (15.4%)	8 (20.5%)	15 (38.5%)	4 (10.3%)	39
C2	Pre-test	6 (15.8%)	6 (15.8%)	5 (13.2%)	5 (13.2%)	7 (18.4%)	9 (23.7%)	38
C2	Post-test	3 (7.9%)	1 (2.6%)	7 (18.4%)	11 (28.9%)	9 (23.7%)	7 (18.4%)	38
				School B				
C3	Pre-test	11 (28.2%)	6 (15.4%)	8 (20.5%)	3 (7.7%)	4 (10.3%)	7 (17.9%)	39
C3	Post-test	3 (8.3%)	5 (13.9%)	6 (16.7%)	8 (22.2%)	8 (22.2%)	6 (16.7%)	36
C4	Pre-test	2 (5.9%)	6 (17.6%)	8 (23.5%)	2 (5.9%)	9 (26.5%)	7 (20.6%)	34
C4	Post-test	2 (5.7%)	5 (14.3%)	13 (37.0%)	0 (0%)	8 (22.9%)	7 (20%)	35

Conclusion about changes in teaching: Comparison groups
One way of summarizing this is to look at changes that one would consider positive in terms of the types of experience we would expect in performance task classes, for example, providing the opportunity for more than one answer, real world situations, etc.

In the following, a positive sign "+" means that for that the change for the class was in the same direction as would support as performance task orientation, a negative sign "-" means it is the opposite direction, and an equal sign "=" means no change (less than a 10% change in percentage). The break down for Questions 15 – 18 is based on whether there is an increase in the percentage for once a week or two weeks or more (the first 3 columns) versus once a month or less (the last 3 columns). For Questions 19 and 20, the breakdown is half or more versus less than a half. This obviously looses some of the distinctions but is useful as a rough guide and provides some basis for comparison. Table 3.32 summarizes the results.

Table 3.32. A Summary about changes in teaching in comparison classes (Questions 15 to 20)

Class	15A	15B	15C	15D	16	17	18	19	20
C1	+	+	-	-	=	-	+	=	+
C2	+	=	=	-	=	-	=	=	+
C3	-	-	-	-	-	-	-	-	+
C4	=	+	=	=	=	-	=	=	=

Note. "+" denotes a positive change towards processes supportive of performance tasks; "-" denotes a negative change towards processes supportive of performance tasks; and "=" denotes no change towards processes supportive of performance tasks.

From the table, it appears that for C3 most of the changes in teaching would not be consistent with changes one might make in a performance task class. For Classes C2 and C4 there is little change while in C1 there are both positive and negative changes in nearly half the items.

Attitude towards performance tasks: Experimental Groups
Part of the post-test questionnaire for the experimental groups asked students questions specifically related to their experience with performance tasks over the intervention period. The following analyses the results on these questions. Table 3.33 is about the results for Question 15 for experimental groups in the surveys.

Table 3.33. Results from Question 15 for experimental classes (I like to solve mathematics questions which have more than more correct answer)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	10 (25.6%)	10 (25.6%)	19 (48.7%)	39
E2	Post-test	8 (21.1%)	11 (28.9%)	19 (50%)	38
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	7 (18%)	8 (20.5%)	24 (61.6%)	39
E4	Post-test	4 (10.5%)	7 (18.4%)	27 (71.1%)	38

The data show that the neighbourhood school is more positive than the high achieving school about solving problems with more than one answer.

Table 3.34 presents the results for Question 16 for experimental groups in the surveys.

Table 3.34. Results from Question 16 for experimental classes (Doing mathematics performance tasks is difficult to me)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	26 (66.6%)	8 (20.5%)	5 (12.8%)	39
E2	Post-test	17 (44.8%)	7 (18.4%)	14 (36.9%)	38
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	17 (43.5%)	6 (15.4%)	16 (41%)	39
E4	Post-test	27 (71.1%)	3 (7.9%)	8 (21.1%)	38

From the table, it appears that there is no obvious difference between the schools. The highest achieving class in the study, E1, as might have been expected, had the smallest percentage of students who found the performance tasks difficult.

Table 3.35 is about the results for Question 17 for experimental groups in the surveys.

Table 3.35. Results from Question 17 for experimental classes (Doing performance tasks helps me to learn mathematics)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	1 (2.6%)	2 (5.1%)	36 (92.3%)	39
E2	Post-test	3 (7.8%)	6 (15.8%)	29 (76.3%)	38
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	6 (15.4%)	2 (5.1%)	31 (79.5%)	39
E4	Post-test	4 (10.5%)	4 (10.5%)	30 (79%)	38

The results on this item show that the vast majority of students felt that doing performance tasks help them learn mathematics, which is very encouraging to the researchers. This positive attitude is important if performance tasks are to become an integral part of the programme. Again the most positive is the best achieving class.

Table 3.36 is about the results for Question 18 for experimental groups in the surveys.

Table 3.36. Results from Question 18 for experimental classes (I like to do mathematics questions which could be solved using different methods)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	2 (5.1%)	2 (5.1%)	35 (89.8%)	39
E2	Post-test	5 (13.1%)	8 (21.1%)	25 (65.8%)	38
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	3 (7.9%)	4 (10.5%)	31 (81.6%)	38
E4	Post-test	3 (7.9%)	2 (5.3%)	33 (86.8%)	38

The results suggest that for three of the four experimental classes over 80% of the students like to solve questions using different methods. For the other class the percentage was lower at 65%. Since performance tasks promote this important objective for mathematics education, this is an encouraging result.

Table 3.37 presents the results for Question 19 for experimental groups in the surveys.

Table 3.37. Results from Question 19 for experimental classes (Doing mathematics performance tasks help me to be more creative in problem solving)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	2 (5.2%)	1 (2.6%)	36 (92.4%)	39
E2	Post-test	5 (13.5%)	8 (21.6%)	24 (64.8%)	37
School B					
Class	Post -test	Disagree	Neutral	Agree	Total
E3	Post-test	2 (5.2%)	4 (10.5%)	32 (84.3%)	38
E4	Post-test	2 (5.4%)	5 (13.5%)	30 (81%)	37

The results show that performance tasks provide an opportunity for creativity. The response to this question indicates that all the experimental classes save this in a positive light.

Table 3.38 is about the results for Question 20 for experimental groups in the surveys.

Table 3.38. Results from Question 20 for experimental classes (I like to solve mathematics questions which have more than more correct answer)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	2 (5.2%)	10 (25.6%)	27 (69.2%)	39
E2	Post-test	6 (16.2%)	6 (16.2%)	25 (67.5%)	37
School B					
Class	Post -test	Disagree	Neutral	Agree	Total
E3	Post-test	1 (2.6%)	7 (18.4%)	30 (78.9%)	38
E4	Post-test	16 (42.1%)	6 (15.8%)	16 (42.1%)	38

The results show that three of the four classes indicate they have to think harder when doing performance tasks, which is what one might expect given the open nature of the tasks. It is not clear why E4 has a quite different pattern.

Table 3.39 is about the results for Question 21 for experimental groups in the surveys.

Table 3.39. Results from Question 21 for experimental classes (I feel lost when I am doing mathematics performance tasks)

		School A			
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	32 (82.1%)	2 (5.1%)	5 (12.9%)	39
E2	Post-test	19 (51.3%)	8 (21.6%)	10 (27%)	37
		School B			
Class	Post -test	Disagree	Neutral	Agree	Total
E3	Post-test	19 (50%)	5 (13.2%)	14 (36.8%)	38
E4	Post-test	27 (71.1%)	2 (5.3%)	9 (23.7%)	38

The figures in the table suggest that most students did not feel lost when they are doing performance tasks, which is quite encouraging. However, we should also note that except for E1, the best class, there are a substantial percentage (more than 20%) of students in all the other classes who felt lost, nevertheless we think this is understandable and consistent with the challenging nature of the tasks. It appears to us that developing students' ability in solving performance tasks is a long term objective for mathematics teachers.

Table 3.40 summarizes the results for Question 22 for experimental groups in the surveys.

Table 3.40. Results from Question 22 for experimental classes (I like to do mathematics questions which involve the real world)

		School A			
Class	Post -test	Disagree	Neutral	Agree	Total
E1	Post-test	2 (5.1%)	0 (0%)	37 (94.9%)	39
E2	Post-test	8 (21.1%)	8 (21.1%)	22 (57.9%)	38
		School B			
Class	Post -test	Disagree	Neutral	Agree	Total
E3	Post-test	4 (10.4%)	6 (15.8%)	28 (73.8%)	38
E4	Post-test	6 (15.7%)	4 (10.5%)	28 (73.7%)	38

Again it is encouraging to see from the table that the majority of the students developed a positive view about working on real-world problems. Nevertheless, it should be also noted that there are clear differences between classes and schools. E1 is very positive about real world problems while, E2, the other class is School A is far less positive. The classes in School B are "in-between".

Table 3.41 summarizes the results for Question 23 for experimental groups in the surveys.

The results presented in the table show a pattern slightly different from Question 22. E1 is still the most positive about the value of doing performance tasks. E2 is more positive about the value than liking working on this type of problem, while E3 and E4 are slightly more negative. Nevertheless, the majority of students in all the classes show positive attitudes.

Table 3.41. Results from Question 23 for experimental classes (Doing mathematics performance tasks helps me see more connections between mathematics and daily life)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	1 (2.6%)	3 (7.7%)	35 (89.7%)	39
E2	Post-test	4 (10.8%)	7 (18.9%)	26 (70.2%)	37
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	3 (7.8%)	9 (23.7%)	26 (68.4%)	38
E4	Post-test	9 (23.6%)	4 (10.5%)	25 (65.8%)	38

Table 3.42 summarizes the results for Question 24 for experimental groups in the surveys.

Table 3.42. Results from Question 24 for experimental classes (Doing mathematics performance tasks helps me to become more systematic when I am solving mathematics problems)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	0 (0%)	1 (2.6%)	38 (97.5%)	39
E2	Post-test	4 (10.8%)	6 (16.2%)	27 (72.9%)	37
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	1 (2.7%)	4 (10.8%)	32 (86.4%)	37
E4	Post-test	3 (7.8%)	5 (13.2%)	30 (78.9%)	38

The results suggest clearly that all the classes feel that doing performance tasks makes them more systematic (with E1 being the most positive) which was an important objective of the performance tasks.

Table 3.43 is about the results for Question 25 for experimental groups in the surveys.

Table 3.43. Results from Question 25 for experimental classes (I need hints to help me do mathematics performance tasks)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	20 (51.2%)	8 (20.5%)	11 (28.3%)	39
E2	Post-test	11 (29.7%)	9 (24.3%)	17 (45.9%)	37
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	10 (25.7%)	6 (15.4%)	23 (59%)	39
E4	Post-test	8 (21.1%)	9 (23.7%)	21 (55.3%)	38

The results shown in the table are quite understandable. As performance tasks are often more challenging, many students still need hints in doing this type of students. Moreover, we can see that E1, the highest achieving class, is different from the other classes. The students in E1 need less hints. However, it is worth noting that about 30% still felt they needed hints. Students in other classes need hints.

Table 3.44 is about the results for Question 26 for experimental groups in the surveys.

Table 3.44. Results from Question 26 for experimental classes (I am good at doing mathematics performance tasks)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	6 (15.3%)	11 (28.2%)	22 (56.4%)	39
E2	Post-test	12 (33.3%)	4 (11.1%)	20 (55.6%)	36
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	10 (25.7%)	8 (20.5%)	21 (53.8%)	39
E4	Post-test	13 (34.3%)	9 (23.7%)	16 (42.1%)	38

The answer to Question 26 is of particular interest. As the performance task data reported earlier showed the level of improvement for most students was not great. In fact, many students declined. The results here suggest that although the students are positive about these tasks, they also realize that there are limitations to their performance. The results imply that more exposure and experience is needed for students to develop more confidence in do this type of challenging problems.

Table 3.45 summarizes the results for Question 27 for experimental groups in the surveys.

Table 3.45. Results from Question 27 for experimental classes (Doing performance tasks takes me more time than doing other mathematics questions usually done in class)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	4 (10.3%)	7 (17.9%)	28 (71.8%)	39
E2	Post-test	10 (27%)	10 (27%)	17 (45.9%)	37
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	7 (18.5%)	12 (31.6%)	19 (50%)	38
E4	Post-test	21 (56.7%)	6 (16.2%)	10 (27%)	37

The results show that the best class, E1, feels that take more time, while in the weakest class, E4, the majority of students feel that the tasks do not take more time. This is probably consistent with the better class being able to do more with the tasks, and understand better the challenging nature of performance tasks.

Table 3.46 summarizes the results for Question 28 for experimental groups in the surveys.

Table 3.46. Results from Question 28 for experimental classes (I would like to have more mathematics performance tasks for my mathematics lessons)

School A					
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	3 (7.7%)	8 (20.5%)	28 (71.8%)	39
E2	Post-test	11 (29.7%)	4 (10.8%)	22 (59.4%)	37
School B					
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	5 (12.8%)	7 (17.9%)	27 (69.3%)	39
E4	Post-test	14 (37.8%)	7 (18.9%)	16 (43.2%)	37

The data show that the response to this question is mixed. The best classes in each school feel they would like more, while the percentage of students who would not like more performance tasks is higher in the two lower achieving experimental classes. The most positive class is the high achieving class, E1. Overall, most students are positive to work more on performance tasks.

Table 3.47 summarizes the results for Question 29 for experimental groups in the surveys.

Table 3.47. Results from Question 28 for experimental classes (Doing mathematics performance tasks makes me learn mathematics better)

		School A			
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	1 (2.6%)	1 (2.6%)	36 (94.8%)	38
E2	Post-test	7 (18.9%)	5 (13.5%)	25 (67.5%)	37
		School B			
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	2 (5.2%)	5 (12.8%)	32 (82.1%)	39
E4	Post-test	4 (10.5%)	3 (7.9%)	31 (81.6%)	38

Again, the results are quite encouraging. In spite of the difficulties some students faced, the vast majority of students in the experimental groups feel that doing performance tasks help them learn mathematics better.

Table 3.48 summarizes the results for Question 30 for experimental groups in the surveys.

Table 3.48. Results from Question 28 for experimental classes (Doing mathematics performance tasks is a waste of time)

		School A			
Class	Post-test	Disagree	Neutral	Agree	Total
E1	Post-test	34 (87.1%)	5 (12.8%)	0 (0%)	39
E2	Post-test	26 (68.4%)	5 (13.2%)	7 (18.4%)	38
		School B			
Class	Post-test	Disagree	Neutral	Agree	Total
E3	Post-test	35 (89.8%)	3 (7.7%)	1 (2.6%)	39
E4	Post-test	25 (65.8%)	5 (13.2%)	8 (21%)	38

The results are somehow striking. The better classes E1 and E3 in each school expressed strong disagreement with the statement that doing mathematics performance tasks is a waste of time, while in the other two classes E2 and E4, about 20% students agree with the statement. Nevertheless, overall most students held positive views doing performance tasks, constituent with their response to other questions as reported earlier.

Conclusion about students' attitude towards performance tasks in experimental groups
A similar approach to summarize the results about students' attitude towards performance tasks in experimental groups can be used to that in the previous section. In the following, the sign "++" means that the perception is generally positive for over 80% based on students' responses to the question concerned, "+" means positive 60 – 79%, "–" means negative for 60 – 69%, "--" means negative for over 80%, while "=" means neutral (neither positive or negative). This is arbitrary but at least means a large proportion of the class feels that way.

Table 3.49 summarizes the results.

Table 3.49. A summary about students' attitude towards mathematics performance tasks in experimental groups (Questions 15 to 30)

Cl	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
E1	=	+	++	++	++	+	++	++	++	++	=	=	+	+	++	++
E2	=	=	+	+	+	+	=	=	+	+	=	=	=	=	+	+
E3	+	=	+	++	++	+	=	+	+	++	=	=	=	+	++	++
E4	+	+	+	++	++	=	+	+	+	+	=	=	=	=	++	+

Note. "++": positive for over 80% of the students; "+": positive for 60%-79% of the students; "-": negative for over 80% of the students; "-": negative for 60%-79% of the students; "=": neutral (neither positive or negative)

From the summary shown in the table, it is clear that there are no negatives. That is, in no cases are 60% or more negative. In fact, only for a couple of questions does the negative reach exactly 50%. The most positive class is E1 from the high achieving school. It has the most "++" and the least number of "=". The next two most positive classes are E3 and E4 from the neighbourhood school while the least positive is E2 from the high achieving school. However, it should be stressed that the overall responses from the students are quite positive about the use of performance tasks.

3.4.4 School-based exam results

As mentioned in Chapter 2, students' regular school-based classroom exams were also used in this sub-project as instruments/benchmarks to measure students' achievement in the cognitive domain during the period of intervention. The data (scores) of students' regular school-based exams were collected from the exams that they took at the end of P2 (Oct/Nov 2003), before the intervention, in mid-P3 (May 2004), at the end of P3 (Oct/Nov 2004), and in mid-P4 (May 2004), which was at the end of the intervention.

In the high-performing school, the following results were obtained for high-performing classes, E1 and C1.

Regarding students' P2 results, the results of t-test show that students in E1 had significantly higher scores than those in C1.

In the mid P3 school exam, which was taken in May 2004, or simply called M2004, students in E1 had significantly higher scores than those in C1; both classes had similar trend (P2→M2004: drop in scores) but the extent of changes in E1 was significantly smaller than that in C1 using statistical procedure GLM (general linear model).

In the end of P3 school exam, which was taken in Oct. 2004 as the final exam of the school year, or simply called F2004, students in E1 had significantly higher scores than those in C1; students in E1 had increase in scores (M2004→F2004, while those in C1 had decrease in scores; moreover, in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM;

In the mid P4 school exam, which was taken in May 2005, the end of the intervention, or simply called M2005, students in E1 had significantly higher scores than those in C1; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM.

Figure 3.1 shows the statistical results in terms of the students' average scores graphically. Note that in the figure, test 1 refers to the end exam of P2 or just P2, while test 2 was taken in May 2004 or M2004, test 3 in Oct. 2004 or F2004, and test 4 in May

2005 or M2005. The same notations apply to other figures about school-based standard exams below.

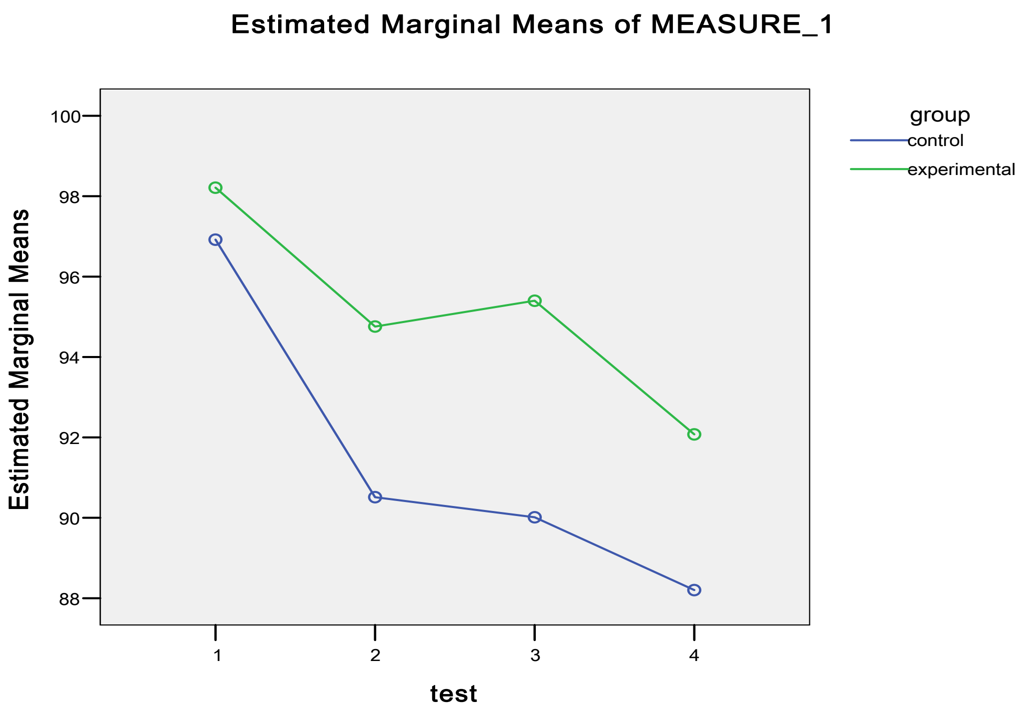


Figure 3.1. School-based standard exam results for classes E1 and C1

For the non-high-performing classes E2 and C2 in the high-performing school, the following results were obtained.

1. Regarding students' P2 results, the results of t-test reveal that students in E2 had higher scores than those in C2, but no significant difference was detected;

2. In M2004 school exam, students in E2 had higher scores than those in C2 but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM;

3. In F2004 exam, students in E2 had lower scores than those in C2 but no significant difference was detected; both classes had similar trend (M2004→F2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM;

4. In M2005 exam, students in E2 had lower scores than those in C2 but no significant difference was detected; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM.

Figure 3.2 shows the statistical results graphically.

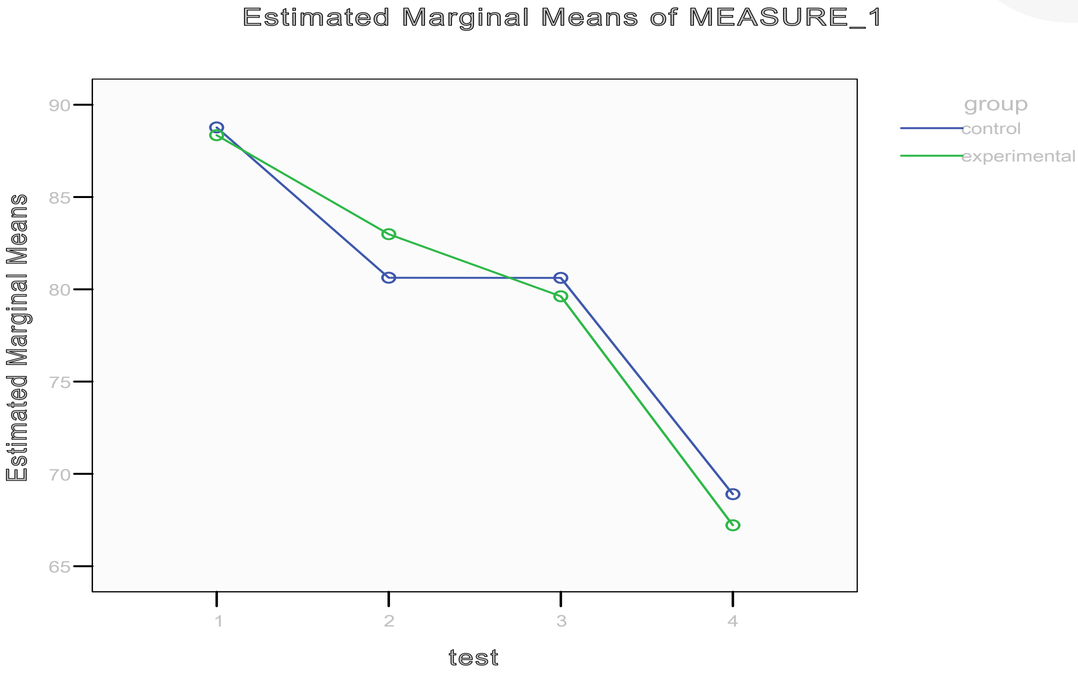


Figure 3.2. School-based standard exam results for classes E2 and C2

In the neighbourhood primary school, regarding the school-based exam scores, we obtained the following results.

1. Regarding students' P2 results, students in both the experimental classes, E3 and E4 had higher scores than those in the comparison classes, C3 and C4, but no significant difference was detected;

2. In M2004, students in the experimental classes had higher scores than those in the comparison classes but no significant difference was detected; both groups had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two groups was detected using statistical procedure GLM;

3. In F2004, students in the experimental classes had lower scores than those in comparison classes but no significant difference was detected; both groups had similar trend (M2004→F2004: increase in scores) and in terms of the extent of changes, no significant difference between the two groups was detected using statistical procedure GLM;

4. In M2005, students in experimental classes had higher scores than those in comparison classes but no significant difference was detected; both groups had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two groups was detected using statistical procedure GLM.

Figure 3.3 shows the statistical results graphically.

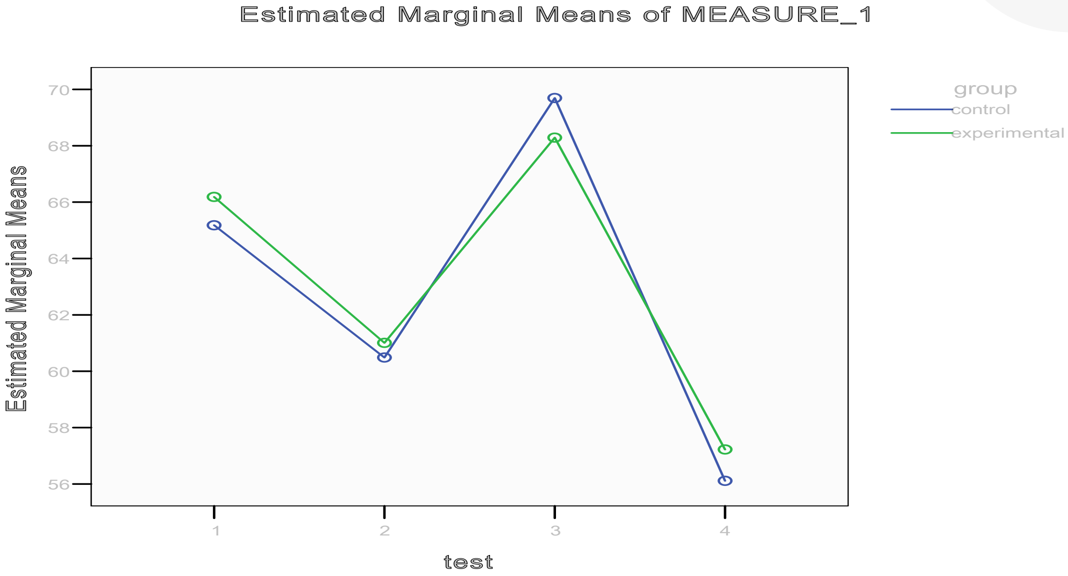


Figure 3.3. School-based standard exam results for classes E2 and C2

Being aware that students' school exam scores are affected by many different factors, we think it appears safe to conclude that the intervention had positive or neutral influences on students' cognitive achievement as measured in standard school exam results.

3.4.5 Student and teacher interview results

Student Interviews
Students were overall positive about their experience in working with performance tasks. Students talked about them being fun, interesting and applicable to real life, etc. The following are excerpts from the interviews:

> Because it's like more fun that you can create things like a graph….. more like real life than the usual textbooks.
>
> It's like you can apply to real life.
>
> Yes, they are interesting.

Students also reported the challenge and thinking involved in these items. They made comments such as:

> The usual ones are easier…because you don't have to think.
>
> Use our brain a lot.
>
> Because if it is too easy, it gets boring.
>
> A bit challenging.
>
> Sometimes I do exams so I need to think.

In looking at these last four comments it is important to emphasis that they reflect things they enjoyed about doing these activities' they like the challenge and having to think.

However, some concern was expressed by the high performing students about the impact of this as it relates to the PSLE. While they felt it was positive there was a concern about the potential negative impact if they did too much. For example, when discussing how often they would like to do this type of activities, one student indicated it should not be too often "*because they do not help with the PSLE.*" It implies that the students might have not fully benefited from the performance tasks because they don't see the integration of this kind of tasks into the high-stake school exams.

Teacher Interviews
The teachers were generally also supportive of the idea. However, they expressed concerns regarding issues such as the time for implementation, and whether is appropriate for all students.

In terms of bring supportive of the idea of using these activities, they said things such as

> T1: It will definitely benefit….it gives the opportunity to stretch my pupils…
>
> *T2: Definitely benefit because they are doing something different.*

When asked to specify in a bit more detail how they felt the tasks benefited their students, some teachers referred to some specific examples. For example:

> T3: I think they discover… I remember the obstacle course… they actually come up with a method of coming up with the answer in a faster way.
>
> T4: Factors… it was something concrete

The teachers also indicated that it gave them insight into their students' work. For example T1 noted that

> …children may be good at computation but they may be weak in other areas. These tasks give me an opportunity …. look at other areas…. It gives me a very strong signal as to their level of math.

As indicated above the general tone was that incorporating tasks was valuable, but there were some reservations regarding who it was valuable for. For example, T2 expressed concerns regarding the ability of the students to respond. She said at different times "*…I think my children won't respond….they will just wait for you to tell them what to do next".* T3 also expressed some concerns regarding the ability of some of her students. She said that, *"The lack of structure is good and not so good at the same time… some students are lost… It's too open for them…"* A similar sentiment was expressed by T4. T2 felt that only about 50% of her students could handle this type of class. T2 also indicated that if a task is beyond them then they find it boring. T3 also indicated that some students might not see the purpose of some tasks.

Teachers have a syllabus to complete. This syllabus is examined through the use of formal examinations at the end of each term. And since these examinations are common across all classes (within a school) in a level the completion of the syllabus is an important consideration. In the discussion with the teachers the issue of time was mentioned. For example, T1 said, "(while I am explaining to) *… the other classes have already moved ahead with the curriculum*." T3 indicated that it takes more time like *any kind of group activity.*

T3 and T4 both mentioned the need to change the assessment. For example T3 suggested that these types of tasks could be made *part of the S2 [Semester 2 assessment], for example.* T4 said "… *you have to change the assessment in order to fit* *this in."* Other comments regarding assessment were mentioned at different times. For example, when discussing future use of activities T2 mentioned that "*the parents still want to see results."*

3.5 Concluding Remarks

In conclusion, given the data and results presented in this chapter, we can conclude the following for each of the research questions

(a) What is the impact of performance tasks on the attitude of students towards mathematics and performance tasks?

In terms of general perception, it would appear that all the classes are not as positive towards mathematics as before, which is not surprising as other researches have also observed the same change of students' attitudes in mathematics from lower grades to higher grades as mathematics become more abstract and challenging. However, the students in experimental groups appeared to have overall more positive attitudes compared to comparison groups. In particular, students in experimental groups showed positive attitude towards performance tasks. In addition, their teachers were also generally positive about performance tasks, though they also reported some practical reservations.

(b) Does project work impact the approach, presentation and ability to solve open-ended performance tasks?

This is a mixed bag. In some cases it improved, other situations it got worse, while in other cases there is no change. For both schools, it would appear that even when there was an overall improvement in the classes performance (based on improvement in level plus no change) in many cases there were many students who performed at a lower level than an higher level of the post-test. The question is how this modifies any conclusion relating to the impact of performance tasks. The research could try and look at specific performance in one category (e.g. approach to the square problem); what was the numerical difference (e.g. how many did better or worse in the presentation on the square problem); when a student changes was it from a 1 to 3 or 2 to 3, etc. (e.g. how many students in the numbers problem improved 1, 2 or 3 levels or went down 1, 2 or 3 levels on solution). However, given the nature of the data this level of data mining does not seem valid. We think that some students' performance in solving performance tasks in the post-test might be related to the fact that students know that their performances in solving these tasks were not taken into account for their final grades, which affected their attitudes and hence behaviours when solving the tasks.

(c) How does solving performance tasks impact performance on traditional semesteral mathematics tests?

The data shows some differences favouring the experimental groups, but is some cases this there was a priori difference in the mathematics performance of the classes. However, it is reasonable to conclude that the use of performance tasks does not appear to have any significantly measurable negative impact on performance on the standard tests used in the schools.

Chapter 4 Results and Findings (II): Student Self-Assessment (Primary)[5]

4.1 Introduction

This chapter reports the component of the MAP which investigated how pupil self-assessment strategy, which is one of the four types of new assessment strategies examined in the project, could be integrated effectively into the mathematics curriculum. Pupil self-assessment strategy was further divided into three sub-types: self-evaluation, self-reflection, and self-constructed tasks.

MAP viewed pupil self-assessment as a means of assessing mainly the affective, but also cognitive, domain of pupil learning and as a desirable life-long learning habit that pupils should be develop. Self-assessment encourages pupils to take responsibilities for their learning, and engages them in reflecting and evaluating their own learning. MAP took the view that Singaporean pupils need support in undertaking more self-directed or self-regulated learning (e.g., see Greene & Azevedo, 2007). The pupils' self-evaluation and self-reflection (i.e., pupils reviewing or looking back at how and what they are learning) should give their teachers access to how the pupils felt and judged their own learning, and enable them to understand their pupils' learning needs better. In addition to the self-evaluation and self-reflection tasks, the pupils undertook the task of constructing and formulating mathematics problems for their classmates to solve.

MAP assumed that the pupils must review and thereby self-assess their mathematics in order to pose and formulate mathematics questions. Such an activity could also help pupils to realise they could be "inventors" of mathematics problems and developed in them views or beliefs favourable toward mathematics. The researchers developed a list of "Ten scenarios & prompts" (details in Appendix 4.1) to help the participating teachers engage their pupils in self-assessment.

Self-assessment is an assessment strategy targeted more at improving pupils' learning than at measuring and grading pupils' performance. In Singapore, such an approach to pupil assessment is in line with educational initiatives such as "Teach Less, Learn More" and "Thinking Schools, Learning Nation". The ability to assess our own learning is a life-skill that is essential for life-long learning. MAP adopted the definition of educational assessment of a pupil as the process of gathering evidence about the pupil's knowledge of, ability to use, and disposition towards mathematics and of making inferences from that evidence for a variety of purposes (NCTM, 1995, p. 3). Pupil self-assessment was defined as a process whereby the teacher gathers information about students' learning through engaging them in reflecting, evaluating and reporting about their learning of mathematics. It yields information in the cognitive and affective domains of the pupil's achievements and performances which pupils may use for improving learning. The teachers, if they have access to the pupil self-assessment data, may utilize them in designing instruction to enhance the mathematics lessons and to inculcate favourable dispositions toward the mathematics. MAP's conceptualization of pupil self-assessment strategy also supports the recommendations of the Singapore Ministry of Education's *Assessment Guides* (MOE 2004a, 2004b). The *Assessment Guides* emphasizes assessment as "an integral component of the teaching and learning process" and the main purpose of assessment is to "improve the teaching and learning of mathematics" (e.g. see MOE 2004a, p.7). In summary, MAP represented an effort to garner research-based evidence on the use of pupil-self assessment in our mathematics classrooms. The design of the Pupil Self-Assessment Sub-Study and its findings are reported in the following sections.

4.2 Research Questions and Conceptual Framework

4.2.1 Research questions

The main research questions for the self-assessment sub-study are:

1. What are the influences of self-assessment strategies on students' learning of mathematics in both cognitive and affective domains?

2. How can the self-assessment strategy be effectively integrated into the mathematics classrooms in Singapore?

In order to be more specific for this subgroup report, the first research question above is further classified into the finding of influences in two domains. They are (a) the cognitive domain and (b) the affective domain. Hence, after the intervention of the self-assessment strategy, the research seeks to find out the following:

1a. Do the pupils improve in their attitudes toward mathematics and the learning of mathematics?

1b. Do the pupils improve in their school mathematics performance after the interventions of self assessment?

4.2.2 Conceptual framework

The focus in self-assessment is to encourage pupils to take greater responsibility for their learning. It is assumed that in doing so, our pupils could be motivated by being placed in the natural flow of learning. The pupils doing self-assessment were asked to evaluate their own learning by using questionnaires crafted by their teachers, by assessing their effort and performance on mathematics tasks set by their teachers, and by posing and constructing mathematics questions for their peers. The teachers then used the survey results, the pupils' self-reports on their work on the mathematics tasks, and the questions constructed by the pupils to address the pupils' needs. Teacher action could be in the form of personal guidance for the pupils individually or redesign of subsequent lessons to address gaps in learning and to develop positive dispositions. Whenever necessary, the teachers used the list of ten prompts to stimulate self-assessment and guide the process along. The self-assessment task which engaged the pupils in constructing mathematics questions to assess their own learning was targeted at enabling the pupils to appreciate mathematics at a different level. It was aimed at avoiding the problem of "teaching to test" which is known to focuses only on what will be tested and neglects other important curricular goals and contents. The pupils assumed the role of "makers of mathematics" when they posed questions or construct problems for their peers. Figure 4.1 below shows the components of the self-assessment sub-study.

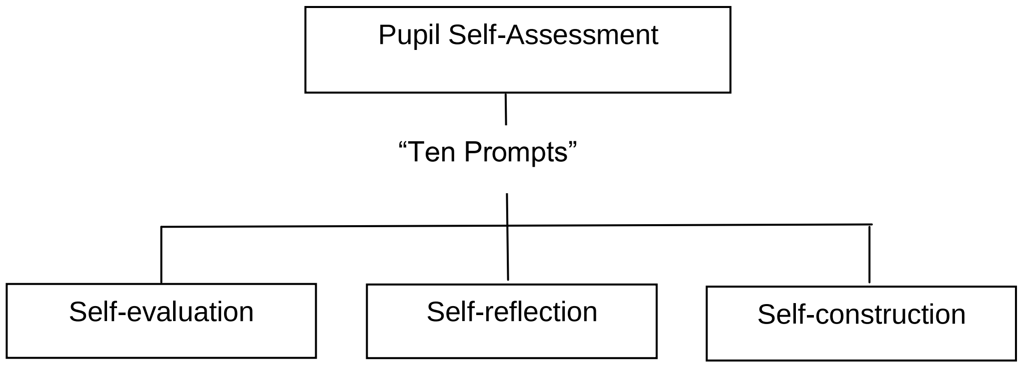


Figure 4.1. A framework for pupil self-assessment

4.3 Methods

4.3.1 Participants

The self-assessment interventions reported in this chapter were carried out in two primary schools: a high performing school and a non-high performing school. A school was ranked as high-performing or non-high-performing based on the how the school's students performed in the PSLE (Primary School Leaving Examinations) and how the schools had helped their pupils in terms of value-addedness. For the study as designed, each school was to nominate four intact Primary 3 classes for the study, to give a total of eight classes from the two schools. The choice of these classes was left to the schools so long as each school identified two of the classes to be *experimental* classes and the other two *comparison* classes. Of the two experimental classes, one was to be a high-performing class and the other a low-performing. There would be equivalent classes for the comparison classes. The schools would also identify the teachers for the sub-study, with the understanding that the teachers consented to participating in the project.

Given that the MAP was a collaborative research effort with schools and given the realities that might be expected in schools, this sub-study was not able to achieve the design requirements for selecting classes and participants. For this sub-study, as it was implemented, each school was able to only provide two experimental classes and one comparison class. The high-performing school considered the three nominated classes as high-performing while the low-performing school considered all three nominated classes' low-performing (See details in Table 4.1).

Pupil participants
As said earlier, two primary schools participated in this sub-study on the use of pupil self-assessment in mathematics. One was a "typical" neighbourhood school (School A) and the other, a high-performing school (School B). Perchance, each school identified three classes (instead of the requested four), two experimental and one comparison. The schools claimed no real differences between the pupils in terms of mathematical ability. The information on these schools is summarized in Table 4.1 below.

Table 4.1. Distribution of classes in self-assessment sub-study

School	Class		
	Experimental	Experimental	Comparison
School A Non-high-performing	Low Ability (LA) Class AE1: Class 3/5	Low Ability (LA) Class AE2: Class 3/7	Low Ability (LA) Class AC1: Class 3/2
School B High-performing	High Ability (HA) Class BE1: Class 3C	High Ability (HA) class BE2: Class 3F	High Ability (HA) Class BC1: Class 3I

Note. (1) Class sizes varied but minimally from year to year since some pupils left at the end of the first year and others joined. However, all classes were about 40 pupils although the number used for analysis was smaller. (2) The respective schools claimed that there were no real differences between the experimental and the comparison classes.

During the 18 months of study, there were some changes to the pupil-participants as the pupils progressed from Primary 3 to Primary 4. In both schools, the change was restricted to pupils being moved from the class to another class, and thereby affecting the profile of the participating class. This sub-study was not abandoned because more than 80% of the pupil-participants remained in their respective classes. That is, in both schools, the number of pupils moved was too small to be considered as a serious threat to the reliability of data that would be gathered and the validity of the inferences that would be made at the end of the sub-study.

Teacher participants
The participating teachers are all trained teachers. As will be noted there is considerable variation in the background and experience of the teachers. It is worth noting that the *Experimental* classes were taught by relatively newly qualified teachers.

Table 4.2. A profile of participating teachers

School A			School B		
Highest qualification and experience			Highest qualification and experience		
AE1	Dip Ed	4 years (same teacher)	BE1	PGDE*	5 years
AE2			BE2	PGDE*	5 years
AC1	PGDE*	14 years	BC1	PGDE*	8 years

Note. * These are postgraduate qualifications following undergraduate degrees.

There was no change in the teacher participants in the low-performing school (School A) for this Sub-study. The same teacher stayed with the project and taught both the *Experimental* classes throughout the intervention period. The teacher for the *Comparison* class followed the class to the next level (P4).

For the high-performing school (School B), one *Experimental* class teacher remained with the project from the start to the end, and taught the same class. The other *Experimental* class teacher left the project at the end of the first school year. Another teacher comparable in qualification (PGDE) and years (7) of teaching experience was assigned to the project at the start of the next year and stay with it to the end. The *Comparison* class teacher stayed with the class for the entire project.

As noted earlier, threats (e.g., changes to participants) to the validity and reliability of such a study in the real-world context of school settings are ever present. The methodological integrity for this sub-study was not seriously challenged because comparable conditions could be re-established to a good extent. Here, MAP was fortunate in that the replacement teacher in School B was comparable in terms of years of teaching. Nevertheless, it would take note of these changes in the analysis and interpretation of the findings.

4.3.2 Instruments for data collection

Questionnaire surveys (Pre- & Post-)
"Pre-tests" were administered to all the 6 classes (4 *Experimental* and 2 *Comparison*) in the two primary schools. The pre-tests consisted of a survey of pupil attitude towards mathematics and three items asking questions related to self-assessment. As mentioned earlier for the entire project, the aim of the pre-tests was to provide baseline information on the pupils' attitude and ability in mathematics. "Post-tests" were conducted at the end of the intervention, about 18 months after the start of MAP. Both the attitude and the three items on self-assessment were designed to be parallel to the pre-tests.

This sub-study used a questionnaire to obtain a measure of the pupils' attitude toward mathematics and to find out about the pupils' experiences in learning mathematics. Part A of the questionnaire contains 22 items asking about the pupils learning attitude toward mathematics. Part B of the survey has 9 items aimed at finding out the ways mathematics lessons had been conducted and how often the pupils had been exposed to the self-assessment task (details of each part in the next two sections). Both Part A and Part B of the survey were administered to all experimental classes and comparison classes before the interventions. However, after the intervention, Part A of the survey was administered to all the experimental classes and the comparison classes and Part B was given to only the comparison classes, as it was not applicable to the experimental classes after intervention.

Part A: Questionnaires on attitude
This section of the questionnaire seeks out the pupils' attitude in terms of their general attitude towards mathematics (G), their beliefs in learning mathematics (B), their feelings

towards their achievement in mathematics (A), and their attitude towards their performance in the learning of mathematics (P). Twelve of the questions are phrased in a positive way while 10 are worded in a negative way to obtain a better measure (better consistency) in the pupils' responses.

The final version of questions about pupils' general attitude (G) in the learning of mathematics is as follows.

Q1: I enjoy doing mathematics.
Q5: Mathematics is hard for me.
Q9: Mathematics is interesting to me.
Q13: I don't have good feelings about mathematics.
Q16: I like spending time on studying mathematics.
Q19: I don't like to attend math lessons.

The final version of questions about students' beliefs (B) in the learning of mathematics is as follows.

Q4: I believe mathematics is useful.
Q8: It is important to know mathematics nowadays.
Q12: Studying mathematics is a waste of time.
Q18: I will use mathematics a lot as an adult.

Below is the final version of questions about students' feeling towards their achievement (A) in the learning of mathematics.

Q2: I am never under a terrible strain in a math class.
Q6: I am not afraid of doing mathematics.
Q10: I am unable to think clearly when doing mathematics.
Q14: I feel lost when trying to solve math problems.
Q17: It makes me nervous to even think about having to do a math problem.
Q20: I have a lot of confidence when it comes to mathematics.

Below is the final version of questions about students' attitude towards their performance (P) in the learning of mathematics.

Q3: I am sure I can learn mathematics well.
Q7: I can get good grades in mathematics.
Q11: I am not good at mathematics.
Q15: I don't think I can do well in mathematics.
Q21: I like solving challenging math problems.
Q22: I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself.

The four main areas of focus of Part A of the survey questionnaire are shown in the Table 4.3.

Table 4.3. Distribution of items in questionnaire

General attitude (G)	Belief (usefulness) (B)	Achievement/Anxiety (A)	Performance (P)
Qn No. 1,5,9,13,16,19 (6 Qns)	Qn No. 4,8,12,18 (4 Qns)	Qn No. 2,6,10,14,17,20 (6 Qns)	Qn No. 3,7,11,15,21,22 (6 Qns)

__Part B: Questionnaire on ways mathematics lessons were conducted__

This part of the questionnaire is designed to enable the researchers to find out how comparable the comparison group is as compared to the experimental group in terms of the new strategy. It aims to find out how often the students were using the alternative assessment and how much the students were aware of the self-assessment. The questions ask about the number of times the pupils were asked to reflect on their learning of mathematics, how often the teacher encouraged alternative solutions to mathematics problems, how often they constructed or posed their own mathematics questions as an instructional strategy, and how often they evaluated their own learning. Though the classification of these survey questions can be subjective, consistency was assured by administering the same questions to the same experimental classes and comparison, before the intervention.

After the intervention, the experimental classes were given a parallel set of items for Part B while the comparison classes answered the same set of questions as the pre-survey.

School-based examinations

As explained earlier, since the pupils took the same PSLE examination, MAP used the results of the PSLE to compare pupil performance across classes and schools. However, the school-based examination was common only for the P3 pupils within the same school. The pupils' performance in the school-based examinations could therefore be used only to compare between the experimental classes and the comparison class within the same school. The MAP tracked the pupils' performance in mathematics in the school-based examinations. Altogether, the researchers collected three sets of examination results, namely Mid-year Examination 2004, End-of-year Examination 2004 and Mid-year Examination 2005. MAP used these examination results to find out whether there was a difference in the mathematics performance of the experimental classes and the comparison class before the start of the intervention.

"New strategy tasks" pre- and post-intervention

The pupils completed a test on self-assessment before and after the intervention period. This self-assessment test was designed to find out how familiar the pupils were with the notion of self-assessment and how well they have developed the habits of self-assessment. Part A of the test consisted of 20 questions on a five-point scale questionnaire. Part B of the test comprised three mathematics questions. Parallel items were used in Part B of both the Pre-test and Post-test. The three test items were written in consultation with the participating teachers to ensure that they would be accessible to their pupils. Item 1 assesses the pupils' ability in mathematical problem solving. Item 2 consists of two parts, (a) and (b), designed to show how well pupils can self-evaluate and apply the same mathematical skill in different situations. Item 2 also serves as a measure of whether pupils in *Experimental* group will perform differently in pre-intervention and post-intervention. Since the pupils in the *Experimental* group have been asked to do self-evaluation and self-reflection during the interventions, it is crucial to MAP to find out whether the intervention has helped the pupils develop the self-assessment skills or apply their mathematical understanding to different situations. In Item 3, the pupils are asked whether they can see the similarity in part (a) and part (b), and give their reasons for it. This question aims to find out whether students are better in reflecting on their learning after intervention and if so, provide the initial confidence for the sub-study to say that the intervention had changed the pupils' predispositions toward learning mathematics.

"New strategy" interventions, data collection, intervention activities

The participating teachers were initially supported by the researchers in the design of the self-evaluation and self-reflection tasks for the topic being taught. The researchers' provided the templates which the teachers modified to suit their pupils and their lessons. They were able to craft their own self-evaluation and self-reflection tasks after the second set of intervention activities without assistance from the researchers. These tasks, in the form of worksheets, focused on finding out whether pupils have difficulties in understanding their lessons and whether they have anything to clarify with their teachers. At the same time the pupils would be helped to consolidate their learning and to continue to reflect about their learning. The worksheets were also crafted to reveal the misconceptions and learning difficulties that the pupils might have. In all, these intervention instruments were aimed at developing the pupils' self-evaluation and self-reflection habits. Samples of the self-evaluation and self-reflection worksheets are shown in the Appendices 4.2, 4.3, 4.4 and 4.5.

The teachers administered the self-evaluation activity at the end of a topic or chapter, and the self-reflection worksheet once a month. During the intervention period, the researchers helped the teachers summarize the evaluations and reflections. The summary though giving the researchers much insight of the intervention, usually did not give timely feedback to the students or the teachers. In fact, teachers made sure that they read through the students reflections before they went for their next lesson. According to the teachers, this was to ensure that they improved on their teaching instructions and fixed the problems that had occurred during the previous lesson.

Classroom observations by researchers
The teacher taking both the experimental classes in the low-performing school agreed to be observed by the researchers when the self-assessment activities would be tried out. The two teachers, one for each of the two *Experimental* classes in the high-performing school declined to be observed. Altogether, the researchers made 6 classroom observations when informed by the participating teacher that the self assessment activities would be carried out. All the participating teachers refused video-recording of the lessons. The classroom observation is an instrument in this project for the researchers to find out how teachers integrate the self-assessment into their mathematics lesson.

Interviews
Interviews with teachers of the *Experimental* classes and their pupils during, and at the end of, the intervention provided further information on how self-assessment could be integrated into the daily mathematics lesson. As said in Chapter 2, the interview questions were broadly put into two main categories. One category was concerned about students' or teachers' own experience and understanding about self assessment. The other category was about pupils' or teachers' opinions or suggestions on the use of self-assessment. Both the pupils and their teachers of the experimental classes were interviewed for their experiences and views on self-assessment.

<u>*Part A: Interview with teachers*</u>
The teachers teaching the *Experimental* classes were interviewed informally during the intervention. They were interviewed with a structured set of questions at the end of the intervention. The goal of the final interview(at the end of the project) was to provide information on their perceptions of performance tasks that would help in terms of any recommendations that might be made regarding using performance tasks within the Singapore system. For all the interview sessions, the researcher spoke with the teacher individually. The general interview questions with the teachers can be found in Appendix 2.4.

For School B, the teacher in the experimental classes took part in the project for its entire duration - 18 months in all. For the informal interviews, the researcher spoke to the teacher after a lesson (in which the researcher was participant observer) during which self-assessment activities were carried out and kept notes on the teacher's views and experiences. For the final interview, the session was audio-taped and transcribed.

For School A, the two teachers were interviewed when the researcher visited the school to collect the self-assessment worksheets. These teachers had declined MAP's requests for a researcher to observe their lessons during which they would conduct the self-assessment activities. One of the teachers took part in the project for the final 18 months, whereas the other teacher who participated in the project since its commencement and for 12 months, before being replaced; the replacement teacher was with the project for only 6 months by the time the project ended. At the end of the project, only the teacher who was with the project for 18 months was available for the interview. The session was audio-taped and transcribed. Due to time constraint and the busy schedule in the school, the other participating school teacher was not interviewed.

Part B: Interview with pupils
At the end of intervention 6 pupils from each experimental class were interviewed using a set of structured interview questions. These pupils were selected by the participating teachers. The researchers set the condition that they would like to interview a sample of students who had participated in the project for 18 months, based on the assumption that such pupils would be better placed to give a more authentic description of their experiences and feelings about the self-assessment tasks. Parallel structured interview questions were used for all the components of the MAP project, with the expectation that each researcher would adapt the questions as appropriate.

The pupils were interviewed in three or four since it was felt that this would be a more supportive environment for the pupils who were still in primary school. They were asked for their feelings about doing the self-assessment activities and the surveys and to suggest ways to improve the integration of the new assessment into classroom teaching. The goal was to provide additional information on their perceptions of self-assessment tasks as well as triangulate the attitude questionnaire data. The outline of the structured interview questions with the students can be found in Appendix 2.5.

Field-notes
The researcher made field-notes during the visits to the schools to observe the class participating in the self-assessment activities. Recall that only the teacher in School A and not School B agreed to the classroom observations.

Others
Many researchers have kept their own diary and notes. All researchers had done their literature review and a comprehensive *Annotated Bibliography on Alternative Assessment in Mathematics* was produced (Fan, Quek, Ng, et al., 2006).

4.3.3 Procedures and data collection

Pupil self-assessment was demonstrated mainly through the following activities designed in this study.

- Pupil self-evaluation of a topic or lesson (topic-based; focus on pupil assessing their learning of a topic)
- Pupil self-reflection at the end of a topic or unit of learning (task-based; focus on pupils monitoring their own efforts in solving a problem)
- Pupil self-constructed task (pupils making up mathematics tasks or problems at the end of a topic or unit of learning).

Pupil self-evaluation
As one of the participating teachers put it “Assessment and evaluation were often inter-related. Assessment refers to the process of gathering information about students’ abilities and using such information to decide on the future instruction. Evaluation, on the other hand, is the process of assigning value to students’ work.” This research focused was on the pupils’ evaluations, where information was collected after lesson to find out what the pupils felt about or how they judged their learning in order to determine some

form of self-assessment by the pupils occurred. A summary of the pupils' self-evaluation is given in Appendix 4.2.

Pupil self-reflection
Pupil self-reflection was conducted as the participating teachers saw fit. It is a less visible and measurable ingredient of self-assessment. The teachers assured the researchers that they used the self-reflection information collected in designing the next lesson for that class of pupils. Sample copies of self-reflection are given in Appendices 4.3 and 4.4.

Pupil self-constructed task
The teachers asked their pupils to construct questions or problems on the topics taught as an instructional strategy to get the pupils to revise those topics. The pupil-constructed questions were edited by their teacher and tried out with other pupils in the same primary level. In the low-performing school, the teacher presented the pupil self-constructed mathematical questions to pupils in a different P3 class. The self-constructed tasks were successfully tried out in the low-performing school; the participating teachers might formulate the questions to while school B managed to get students to set questions but did not use the students' questions for their class test. Sample copies of self-constructed tests for Express and Normal classes with pseudo names are given in Appendix 4.5.

Ten Prompts and Instructional Approach
"Ten Prompts" is a set of questions and suggestions our MAP researchers crafted to prompt teachers and pupils to deeper thinking and reflection in their teaching and learning. The prompts are attached in Appendix 4.1, as mentioned.

The aim of the intervention was for the teachers in the experimental classes to integrate self-assessment into their regular teaching so that the pupils would see self-assessment tasks as an integral component of instruction, not as an add-on to regular classroom instruction and assessment.

4.3.4 Limitations of the study

For a study in a school-based setting that stretched over about 18 months, it might be expected that there would be changes to the profile of the participating teachers and pupils. These changes, however minimal in this sub-study, could undermine the dependability of the data collected over the period of intervention.

Limitations of study in School A
The change of pupil participants in lower-performing School A was restricted to movement of three pupils to another class. However, there was no change in the teacher-participants for both the *Experimental* and *Comparison* classes in this school. Thus, in terms of participant changes (i.e., mortality), the data from School could be said to be dependability. MAP recognises, however, that changes no matter how minimal might still impinge on the validity of inferences drawn and the reliability estimates of the intervention data collected from the pupils. It therefore remains cautious when drawing inferences based on the intervention data collected from School A.

Limitation of study in School B
As described earlier, there were changes to the teacher-participants and the pupil-participants in School B. Although the form teacher of one of the *Experimental* classes had to be replaced at the end of 12 months, MAP was fortunate that the replacement was comparable in terms of qualifications (but with more years of teaching). Another unanticipated development was that the replacement teacher went on medical leave and her *Experimental* class was taken by a relief teacher for two months. Still MAP was able to persuade the relief teacher and the replacement teacher to continue with the interventions. Pupil changes were due to transfers from one class to another for reasons

undisclosed by the school. From the figures provided by the participating class teachers, an estimated 80% of the original pupils remained with this sub-study. As such MAP acknowledges a serious threat to validity posed by the changes in participants in School B.

4.4 Results

We report the findings of this sub-study in the following sections, beginning with the results from questionnaire surveys.

4.4.1 Results from questionnaire

The results are reported in three sub-sections, namely, attitudes towards mathematics and learning of mathematics, items on ways mathematics lessons were conducted, and attitudes towards self-assessment strategies.

Attitudes towards mathematics and learning of mathematics
We start with School A. As said earlier, School A is a typical neighborhood school.

Comparing within classes in School A
Table 4.4 provides a summary of the general results in School A on both Part A and Part B of the pre- and post- survey in the study.

Table 4.4. Results of pre-survey and post-survey for School A

	Pre-Survey		Post-Survey	
	Part A	Part B	Part A	Part B
Experimental/Low (EL): Class 5	P	L (F:19)	P (N:13)	P
Experimental/Low (EL): Class 7	P	L (F:15b, 19)	P	P
Comparison/Low (CL): Class 2	P (N:7, 10)	L (F:15a, 19)	P (N:7, 10, 11, 13)	L (F:19)

Note. “P” indicates positive attitude (average > 5), otherwise N, followed by the question number. “L” less frequent (average >= 3), otherwise F, followed by the question number.

The survey results also show that students in *Experimental* classes in School A provided more positive responses in the post-survey than pre-survey on all but 6 items (Items 2, 3, 4, 8, 13, and 14); and on 1 item (Item 10) they provided significantly more positive responses and on 1 item (Item 4) they provided significantly more negative responses in the post-survey.

On the other hand, students in *Comparison* classes provided more negative responses in the post-survey than pre-survey on all but 1 item (Item 8) and on 3 items (Items 6, 13, and 14) they provided significantly more negative responses in the post-survey.

Table 4.5 provides a summary of the results about the change of students’ attitudes in School A from the pre-survey to the post- survey in the study.

In both the pre-survey and post-survey, students in *Experimental* classes provided more positive responses than those in *Comparison* classes; in particular, it was the case for 8 items in the pre-survey and 13 items in the post-survey; moreover, while no significant difference between the two types of classes was detected in the pre-survey, students in *Experimental* classes provided significantly more positive responses on 6 items (Items 1, 9, 10, 11, 12, and 14) in the post-survey.

Table 4.5. Change of students' attitudes in school a from pre-survey to post- survey

	Overall Trend in Part A	Overall Trend in Part B
EL, Class 4/5	P Ex:2, 3, 4, 9, 10, 13, 14 No Sig	
EL, Class 4/7	P Ex:2, 3, 4, 8, 13, 14 Sig:10	
CL, Class 4/2	N Ex:8 Sig:6, 13, 14	L (F:15d, 18) Sig:19

Note. (1) "N" indicates compared to pre-survey, students' attitude decreased, otherwise "P". (2) "Ex" indicates exceptional cases followed by question number.

A summary of the comparison results in School A on both Part A and Part B of the pre-survey and post-survey in the study is provided in Table 4.6(a) and Table 4.6(b), respectively.

Table 4.6(a). A comparison of the results between classes in School A on both Part A and Part B of pre-survey

	Part A		Part B	
	EL	CL	EL	CL
EL	Class 3/5 > 3/7 Ex:1, 7, 13 Sig:10	EL > CL Ex:1, 4, 5, 6, 9, 14 No Sig	Class 3/5 < 3/7 Ex:15b, 17, 20 No Sig	EL > CL Ex:15b, 15d, 17, 20 Sig:15a
CL	EL > CL Ex:1, 4, 5, 6, 9, 14 No Sig	-	EL > CL Ex:15b, 15d, 17, 20 Sig:15a	-

Note. ">" indicates more positive; "Ex" indicates exceptional cases; "Sig" indicates that there is significant difference followed by question number.

Table 4.6(b). A comparison of the results between classes in School A on both Part A and Part B of post-survey

	Part A		Part B	
	EL	CL	EL	CL
EL	Class 4/5 > 4/7 Ex:9, 10, 11, 13 Sig:4	EL > CL Ex:8 Sig:1, 9, 10, 11, 12, 14	Class 4/5 > 4/7 Ex:15, 16, 17, 22, 24, 31, 32, 34, 35 (No Sig)	
CL	EL > CL Ex:8 Sig:1, 9, 10, 11, 12, 14	-		

Note. ">" indicates more positive; "Ex" indicates exceptional cases; "Sig" indicates that there is significant difference followed by question number.

Comparing within classes in School B
Below we report the results in School B, which is a high-performing school. Table 4.7 provides a summary of the general results in School B on both Part A and Part B of the pre- and post- survey in the study.

Table 4.7. Results of pre-survey and post-survey for School B

	Pre-Survey		Post-Survey	
	Part A	Part B	Part A	Part B
Experimental/High (EH) Class C	P (N:13)	L (F:19)	P	P (N:26, 27, 32)
Control/High (CH) Class I	P	L (F:16, 19)	P	L (F:19)
Experimental/High (EH) Class F	P	L (F:19)	P	N (P:17,18, 24, 31)

Note. “P” indicates positive attitude (average > 5), otherwise N, followed by the question number. “L” less frequent (average >= 3), otherwise F, followed by the question number.

The results show that pupils provided positive responses to all but 10 items (Item15, Item19, Item20, Item22, Item26, Item27, Item28, Item29, Item32, Item33).

The results also reveal that pupils in *Experimental* classes provided more negative responses in the post- than pre-survey on all but 2 items (Item10, Item13) and on 2 item (Item3, Item12) they provided significantly more negative responses in the post-survey;

Pupils in *Comparison* classes provided more negative responses in the post- than pre-survey on all items and on 5 items (Item 2, Item 3, Item 4, Item 5, Item 14) they provided significantly more negative responses in the post-survey;

Table 4.8 provides a summary of the results about the change of students’ attitudes in School B from the pre-survey to the post- survey in the study.

Table 4.8. Change of students’ attitudes in School B from pre-survey to post- survey

	Overall Trend in Part A	Overall Trend in Part B
EH, Class C	N Ex: 1, 7, 10, 13, 14. No Sig	
CH, Class I	N Sig:2, 3, 4, 5, 14	L (F: 15b, 20). Sig:15c, 20
EH, Class F	N Ex: 13. Sig:12	

Note. (1) “N” indicates compared to pre-survey, students’ attitude decreased, otherwise “P”. (2) “Ex” indicates exceptional cases followed by question number.

The results of comparison in School B on both Part A and Part B of the pre-survey and post-survey in the study are provided in Table 4.9(a) and Table 9(b), respectively.

In the pre-survey, pupils in *Experimental* classes provided more negative responses than those in *Comparison* classes on all but 3 items (Item 7, Item 9, Item 12); in the post-survey, they provided more positive responses on all but 4 items (Item 3, Item 6, Item 8, Item 12); moreover, no significant difference between the two types of classes was detected in both the pre- and post-survey.

Table 4.9(a). A comparison of the results between classes in School A in both Part A and Part B of pre-survey

	Part A		Part B	
	EH	CH	EH	CH
EH	3C < 3H Ex:11 No Sig	EH < CH Ex:7, 9, 12 No Sig	3C < 3F Ex:16, 18, 20 Sig:15a	EH > CH Ex:17, 20 Sig:15a, 15c, 15d
CH	EH < CH Ex:7, 9, 12 No Sig	-	EH > CH Ex:17, 20 Sig:15a, 15c, 15d	

Note. ">" indicates more positive; "Ex" indicates exceptional cases; "Sig" indicates that there is significant difference followed by question number.

Table 4.9(b). A comparison of the results between classes in School B on both Part A and Part B of pre-survey

	Part A		Part B	
	EH	CH	EH	CH
EH	4C < 4F Ex:1, 4, 9, 11, 12, 14 No Sig	EH > CH Ex:3, 6, 8, 12 No Sig	4C > 4F No Sig	
CH	EH > CH Ex:3, 6, 8, 12 No Sig			

Note. ">" indicates more positive; "Ex" indicates exceptional cases; "Sig" indicates that there is significant difference followed by question number.

Questionnaires on ways mathematics lessons were conducted.
For the survey on the ways mathematics lesson were conducted, the mode was used as the measure of central tendency.

The survey results shed light on how the *Experimental* classes and the *Comparison* classes differed on in terms of the new assessment strategies.

Pre-survey results for both schools
For all participating classes (*Experimental* and *Comparison*) in both School A and School B, the survey results revealed that mathematics lessons were conducted without much use of the new strategies except for the items "teacher asks pupils to have more than one correct answer".

For illustration purposes, Table 4.10(a) presents the pre-survey results in Experimental Class 3C in School B (high-performing school).

Table 4.10(a). Pre-survey results about how Mathematics lessons were conducted in School B (Experimental Class 3C)

Qn15 – Qn18	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total	Average
Q15A: my math teacher had asked me to write down the reasons for my math answers	6	3	0	1	4	22	36	4.67
	16.7%	8.3%	0.0%	2.8%	11.1%	61.1%		
Q15B: my math teacher had asked me to explain mathematics to the whole class.	8	3	2	3	4	16	36	4.11
	22.2%	8.3%	5.6%	8.3%	11.1%	44.4%		
Q15C: my math teacher had asked me to write down my feelings about mathematics.	0	2	1	0	0	34	37	5.70
	0.0%	5.4%	2.7%	0.0%	0.0%	91.9%		
Q15D: my math teacher had asked me to explain math ideas in writing.	3	2	2	3	3	24	37	4.97
	8.1%	5.4%	5.4%	8.1%	8.1%	64.9%		
Q16: my math teacher encouraged me to solve math questions in different ways.	2	10	5	2	5	12	36	3.94
	5.6%	27.8%	13.9%	5.6%	13.9%	33.3%		
Q17: my math teacher asked me to make up math questions by myself.	1	3	11	7	6	8	36	4.06
	2.8%	8.3%	30.6%	19.4%	16.7%	22.2%		
Q18: how often did your math teacher ask you think about the reason for your solving math problems?	5	7	5	2	4	14	37	3.95
	13.5%	18.9%	13.5%	5.4%	10.8%	37.8%		
Qn19 – Qn20	Almost all	More than half	Half	Less than half	A Few	None	Total	Average
Q19: how many math questions did your teacher ask you to do have more than 1 correct answer?	17	9	2	1	5	3	37	2.38
	45.9%	24.3%	5.4%	2.7%	13.5%	8.1%		
Q20: how many math questions did your teacher ask you to do have nothing to do with real life situations?	5	5	4	6	10	8	38	3.92
	13.2%	13.2%	10.5%	15.8%	26.3%	21.1%		

Table 4.10(b) presents the pre-survey results in Experimental Class 3F in School B.

Table 4.10(b). Pre-survey results about how Mathematics lessons were conducted in School B (Experimental Class 3F)

Qn15 – Qn18		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total	Average
Q15A: my math teacher had asked me to write down the reasons for my math answers	n	0	0	4	0	0	34	38	5.68
	%	0.0%	0.0%	10.5%	0.0%	0.0%	89.5%		
Q15B: my math teacher had asked me to explain mathematics to the whole class.	n	3	7	4	3	3	18	38	4.32
	%	7.9%	18.4%	10.5%	7.9%	7.9%	47.4%		
Q15C: my math teacher had asked me to write down my feelings about mathematics.	n	0	1	1	1	0	35	38	5.76
	%	0.0%	2.6%	2.6%	2.6%	0.0%	92.1%		
Q15D: my math teacher had asked me to explain math ideas in writing.	n	1	2	3	0	3	29	38	5.34
	%	2.6%	5.3%	7.9%	0.0%	7.9%	76.3%		
Q16: my math teacher encouraged me to solve math questions in different ways.	n	9	6	9	2	3	9	38	3.29
	%	23.7%	15.8%	23.7%	5.3%	7.9%	23.7%		
Q17: my math teacher asked me to make up math questions by myself.	n	2	5	8	3	4	16	38	4.32
	%	5.3%	13.2%	21.1%	7.9%	10.5%	42.1%		
Q18: how often did your math teacher ask you think about the reason for your solving math problems?	n	7	5	6	3	5	12	38	3.79
	%	18.4%	13.2%	15.8%	7.9%	13.2%	31.6%		
Qn19 – Qn20		Almost all	More than half	Half	Less than half	A Few	None	Total	Average
Q19: how many math questions did your teacher ask you to do have more than 1 correct answer?	n	16	11	1	1	8	1	38	2.39
	%	42.1%	28.9%	2.6%	2.6%	21.1%	2.6%		
Q20: how many math questions did your teacher ask you to do have nothing to do with real life situations?	n	8	2	7	3	11	7	38	3.74
	%	21.1%	5.3%	18.4%	7.9%	28.9%	18.4%		

Table 4.10(c) presents the pre-survey results in Comparison Class 3I in School B. By the way, no significant difference was found between the two schools.

Table 4.10(c). Pre-survey results about how Mathematics lessons were conducted in School B (Comparison Class 3I)

Qn15 – Qn18	Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never	Total	Average
Q15A: my math teacher had asked me to write down the reasons for my math answers	9	5	4	3	3	13	37	3.68
	24.3%	13.5%	10.8%	8.1%	8.1%	35.1%		
Q15B: my math teacher had asked me to explain mathematics to the whole class.	7	4	5	1	4	16	37	4.05
	18.9%	10.8%	13.5%	2.7%	10.8%	43.2%		
Q15C: my math teacher had asked me to write down my feelings about mathematics.	2	4	3	3	5	20	37	4.76
	5.4%	10.8%	8.1%	8.1%	13.5%	54.1%		
Q15D: my math teacher had asked me to explain math ideas in writing.	7	7	5	0	4	13	36	3.72
	19.4%	19.4%	13.9%	0.0%	11.1%	36.1%		
Q16: my math teacher encouraged me to solve math questions in different ways.	10	10	3	3	3	7	36	3.00
	27.8%	27.8%	8.3%	8.3%	8.3%	19.4%		
Q17: my math teacher asked me to make up math questions by myself.	1	5	7	1	6	17	37	4.54
	2.7%	13.5%	18.9%	2.7%	16.2%	45.9%		
Q18: how often did your math teacher ask you think about the reason for your solving math problems?	9	6	7	6	1	8	37	3.22
	24.3%	16.2%	18.9%	16.2%	2.7%	21.6%		
Qn19 – Qn20	Almost all	More than half	Half	Less than half	A Few	None	Total	Average
Q19: how many math questions did your teacher ask you to do have more than 1 correct answer?	15	12	5	1	3	2	38	2.24
	39.5%	31.6%	13.2%	2.6%	7.9%	5.3%		
Q20: how many math questions did your teacher ask you to do have nothing to do with real life situations?	3	3	5	3	11	13	38	4.45
	7.9%	7.9%	13.2%	7.9%	28.9%	34.2%		

Post-survey results for both schools

The post-survey results also suggested an increased awareness of self-assessment strategies in both schools. While we would like to attribute the heightened awareness to the fact that the pupils in the experimental classes now have had experience with self-assessment, we would also not rule out that the awareness generated could simply be due to the fact that the teacher mentioned the names of the strategies during those lessons where the interventions were carried out. However, comments from the participating teachers during the interviews suggested that the pupils were more aware of self-assessment because of their direct experience with self-assessment activities and not merely the mention of the terms ‘self-assessment,’ or ‘self-reflection’ or ‘self-constructed task.’

Attitudes towards self-assessment strategies

As explained earlier, students’ attitudes towards self-assessment is only conducted in the post-survey with the experimental classes. It contains 21 items in Part B of the survey (note that students’ responses to some items have reported earlier). The results revealed that students provided positive responses to all the items, and no significant difference was detected between the experimental classes.

Overall, the patterns of responses indicate that the pupils were generally positive (agreeing), or they are not adverse (disagreeing), about the value and use of self-assessment. For instance, their responses suggest that they were ambivalent about their liking for mathematics self-reflection (Qn 15) and about doing more self-evaluation (Qn 19; Qn 22), although doing so made them (49% Agree, 27% Neutral and 24% Disagree) think of what they had learnt (Qn 16). Slightly more than half (54%) disagreed and about a fifth (19%) agreed that self-evaluation as being a waste of time (Qn 17); the rest being neutral about it. In addition, slightly more than half (58%) felt that the self-evaluation questions made them more aware of their areas of weaknesses (Qn 18) and about the same number (54%) thought the self-evaluation made them learn better (Qn 21).

The following are some favourable responses to the use of self-evaluation:

- "doing mathematics self-reflection questions make me look back at what I have done (Qn 23), with 62% agreeing and 16% disagreeing;
- "doing self-reflection questions help me spot my own mistakes when doing mathematics (Qn 25), with 57% agreeing and 22% disagreeing; and
- "creating mathematics on my own is a waste of time (Qn 31), with 65% disagreeing and 14% agreeing.

Table 4.13 provides the statistical results in terms of percentage in Experimental Class 4F for the 21 items. As said earlier, no significant difference was detected between the experimental classes.

Table 4.13. Students' responses to 21 items about self-assessment strategies in post-survey (Experimental Class 4F)

	Disagree	Neutral	Agree	Total
Q15: I like to do mathematics self-evaluation.	14	8	11	33
	0.425	24.20%	0.333	
Q16: Doing mathematics self-evaluation questions helps me to think about what I have learnt.	13	4	16	33
	0.393	12.10%	0.484	
Q17: Doing mathematics self-evaluation questions is a waste of time.	19	4	10	33
	0.576	12.10%	0.303	
Q18: Doing mathematics self-evaluation questions makes me more aware of where I am weak in.	14	2	17	33
	0.425	6.10%	0.516	
Q19: I like to do more self-evaluation for my mathematics lessons.	16	7	10	33
	0.485	21.20%	0.304	
Q20: Doing self-evaluation is an important part of learning mathematics.	15	7	11	33
	0.455	21.20%	0.334	
Q21: Doing self-evaluation questions makes me learn mathematics better.	19	3	11	33
	0.576	9.10%	0.333	
Q22: I like to do mathematics self-reflection.	21	3	9	33
	0.637	9.10%	0.272	
Q23: Doing mathematics self-reflection questions make me look back what I have done.	13	6	14	33
	0.394	18.20%	0.425	
Q24: Doing mathematics self-reflection is a waste of time.	19	2	12	33
	0.576	6.10%	0.364	

Q25: Doing self-reflection questions helps me to spot my own mistakes when doing mathematics.	11	6	16	33
	0.334	18.20%	0.486	
Q26: Doing mathematics self-reflection questions helps me to learn new things.	13	6	14	33
	0.394	18.20%	0.425	
Q27: I like to do more self-reflection for my mathematics lessons.	18	6	9	33
	0.545	18.20%	0.273	
Q28: Doing self-reflection makes me learn mathematics better.	17	4	12	33
	0.515	12.10%	0.364	
Q29: Doing self-reflection is an important part of learning mathematics.	16	5	12	33
	0.485	15.20%	0.365	
Q30: I enjoy creating my own mathematics questions.	13	5	15	33
	0.394	15.20%	0.455	
Q31: Creating mathematics on my own is a waste of time.	17	7	9	33
	0.515	21.20%	0.272	
Q32: I am good at creating mathematics questions	17	6	10	33
	0.515	18.20%	0.303	
Q33: I like to create more questions by myself for my mathematics lessons.	15	8	10	33
	0.454	24.20%	0.303	
Q34: Creating my own mathematics questions makes me learn mathematics better.	11	8	14	33
	0.332	24.20%	0.425	
Q35: Creating my own mathematics questions is an important part of learning mathematics.	13	6	14	33
	0.394	18.20%	0.424	

4.4.2 Results from school-based examinations

As it is indicated in Chapter 2, comparison within classes and between classes for both the schools were carried out using the results from school-based mathematics examination and the pre-intervention and post-intervention surveys.

School-based examination results for School A
The following results were obtained for School A, the neighborhood school participating in this sub-project.

1. Regarding students' P2 results, t-test results show that students in experimental classes had higher scores than those in comparison class but no significant difference was detected.

2. In M2004 school exams, t-test results show that students in experimental classes had higher scores than those in comparison class but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using GLM.

3. In F2004 school exams, students in experimental classes had lower scores than those in comparison class but no significant difference was detected using t-test; both classes had similar trend (M2004→F2004: increase in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using GLM.

4. In M2005 school exams, students in experimental classes had lower scores than those in comparison class but no significant difference was detected using t-test; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using GLM.

Figure 4.1 presents the statistical results in terms of the students' average scores graphically. Again, in the figure, test 1 refers to the end exam of P2 or just P2, while test 2 was taken in May 2004 or M2004, test 3 in Oct. 2004 or F2004, and test 4 in May 2005 or M2005. The same notations apply to other figures about school-based standard exams below.

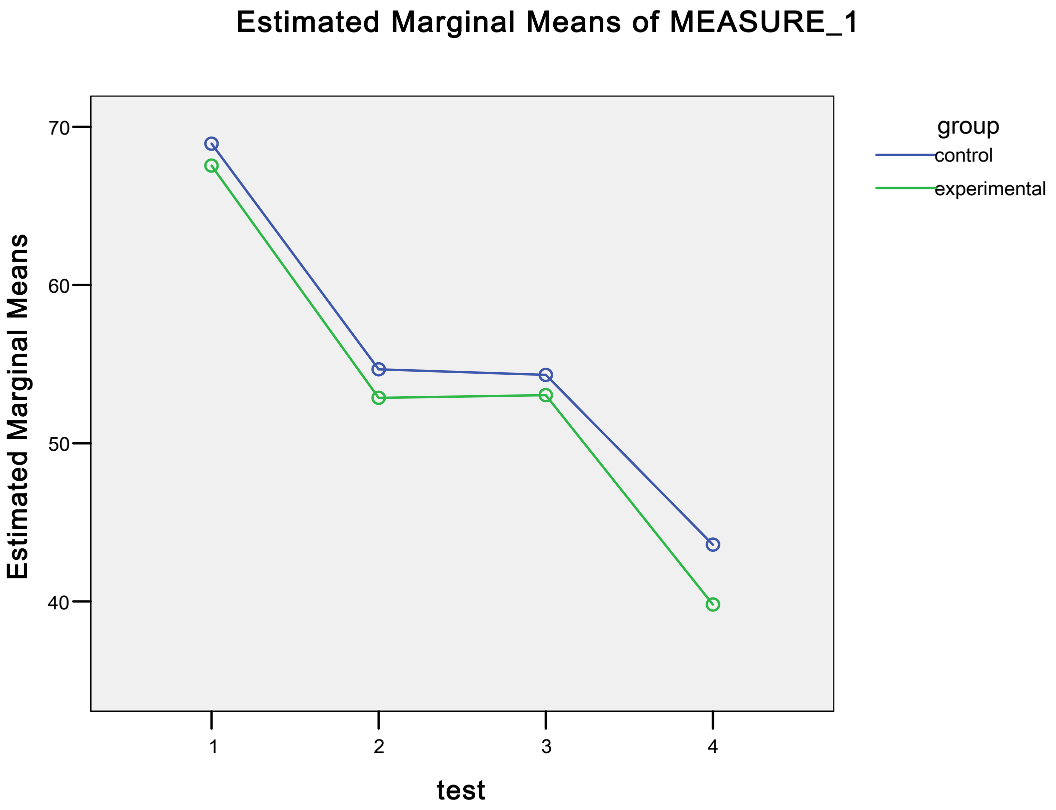


Figure 4.1. School-based standard exam results for experimental and comparison classes in School A

School-based examination results for School B
The results of students' performance in School B, the high-performing school, in school-based examinations are presented below.

1. Regarding students' P2 results, t-test results show that students in *Experimental* classes had significantly higher scores than those in *Comparison* class.

2. In M2004 school exams, students in *Experimental* classes had higher scores than those in *Comparison* class but no significant difference was detected; both classes had similar trend (P2→M2004: increase in scores) and in terms of the extent of changes, no significant difference between the two classes was detected according to GLM.

3. In F2004 school exams, t-test results reveal that students in *Experimental* classes had significantly higher scores than those in *Comparison* class but no significant difference was detected; both classes had similar trend (M2004→F2004: drop in scores) but the extent of changes in *Experimental* classes was significantly smaller than that in *Comparison* class [GLM].

4. In M2005 school exams, t-test results again suggest that students in *Experimental* classes had significantly higher scores than those in *Comparison* class but no significant difference was detected; students in *Experimental* classes had decreased in scores (F2004→M2005), while those in *Comparison* class had increase in scores; moreover, in terms of the extent of changes, no significant difference between the two classes was detected using GLM.

Figure 4.2 shows the statistical results in terms of the students' average scores graphically in School B.

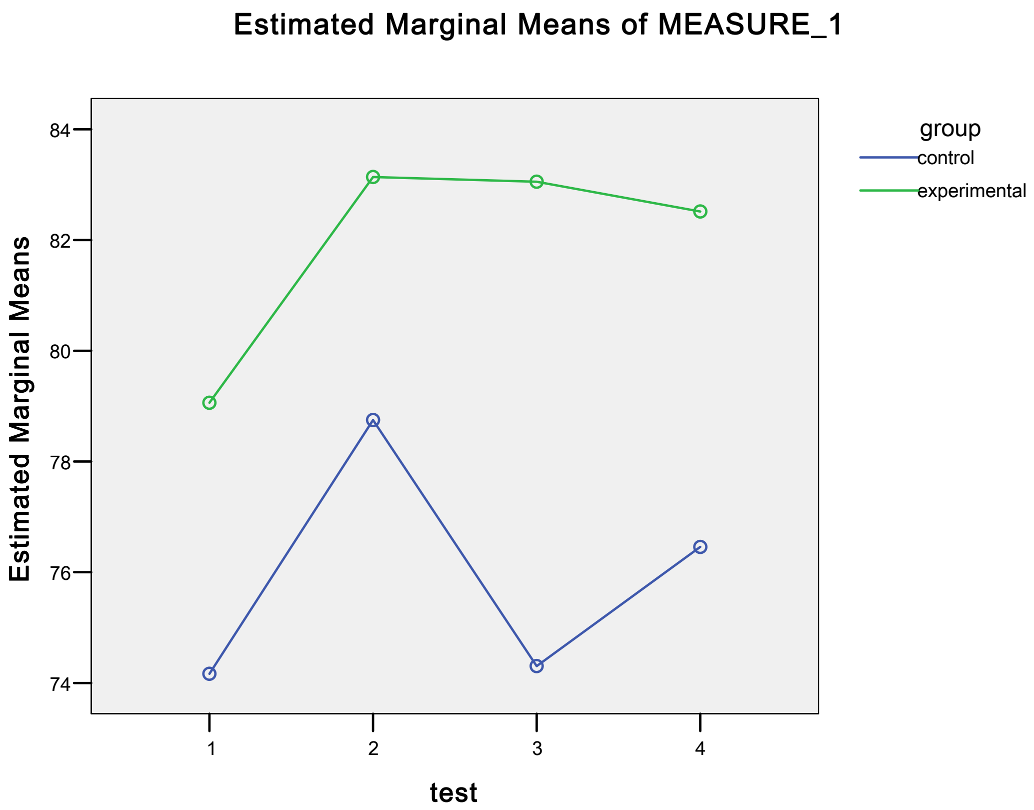


Figure 4.1. School-based standard exam results for experimental and comparison classes in School A

In summary, given the statistical results, we think the intervention had produced neutral influences on students' performance on school-based examination. Again, it should be pointed out that students' school exam scores are a function of many different variables.

4.4.3 Results from interviews

Interview with teachers
The interviews with teachers revealed that all the participating teachers, speaking from their perspective as mathematics teachers, were very supportive of incorporating self-assessment in their teaching. All of the teachers interviewed reported that they got to know their pupils better through use of self-assessment strategies. They would also use self-assessment activities now and then as variation to their normal classroom routines.

The interviews also found that the teachers in the experimental classes were very confident in the integrating self assessment into their classroom teaching. As classroom teachers they are supportive of the self-assessment activities in that these activities gave them additional information about their pupils' learning. All the teachers believed self-assessment to be a good instructional strategy that will benefit in this teaching and learning of mathematics.

Nevertheless, they found difficulty using them as frequently as suggested by the MAP. Unless stipulated as a departmental or school policy, they felt that there would be difficulty in carving time from the packed, as well as lock-step, curriculum. A self-assessment activity would take 10 to 15 minutes to complete (start to end, including collection of the self-assessment sheets). They found it challenging to shave time off from doing the prescribed set of exercises to carry out the self-assessment activities. The pupils found them a novelty initially but took to the tasks later as they would any other classroom activities (even when a researcher was present in the classroom).

The participating teacher in School A remained positive throughout the intervention period. She wrote that the pupils were very excited about the tasks but had great difficulties in answering (in writing) the self-reflection questions, which she attributed to their poor understanding of the questions. Thereafter, she phrased the questions for all subsequent self-assessment tasks very directly and in very simple terms.

As the intervention unfolded, the teacher reported difficulty in finding time to read through the pupils' reflection. Also, the pupils were finding it "routine" to have to answer similar type of questions for the self-assessment and self-reflection activities. A telling comment from the teacher was:

> many of them [the pupils] knew they were careless in their tasks [for the self-reflection activity] were still careless in doing their Maths sum.

We infer from her remark that there is more to be done in helping the pupils work carefully on Maths sums even though they might have become aware of their own careless way through the self-assessment activities.

The teacher was also concerned about the ability of the pupils to self-assess. She observed that the performance of the pupils in tests remained poor although they "indicated in the topical reflection questionnaire that they understood and were able to do the sums." At face value, therefore, it would appear that the pupils were unable to accurately self-assess but, then again, there could be other intervening or mediating factors that caused the pupils to do poorly in the tests.

In School B (higher-performing school), teachers found the self-assessment useful in providing them with additional information on the pupils' learning in a more organised way. This, as one of them carefully put it, was not to say that they did not seek feedback on their teaching and pupil learning, but that they would otherwise rely on their informal observations, gut feelings and oral questioning to do so.

Both teachers in the school felt that self-reflection might be beyond the pupils at Primary 3 level. The pupils were able to talk but not write about their experience and thinking. The teachers felt that the self-reflection tasks (to quote one of the teachers) "would certainly be more useful for children who had reached a higher level of maturity whereby they are more determined to find out why they aren't faring better."

Interviews with pupils

During the experimental period, teachers in the experimental class use evaluation and reflection worksheets to find out their pupils' learning difficulties and use the findings to improve on their teaching. The pupils were also asked to construct questions on the topics they learnt.

The researchers interviewed three pupils per experimental class. The interviews found that through self evaluation, they were made aware of their mistakes but interventions of other kind might be needed to help them correct their mistakes and improve the performance on the tests.

In addition, in the chat with the interviewer, the pupils reported the self-constructed tasks to be the most fun. To them, the tasks were seen as being "different" from normal type of activities (presumably, the typical explanation-illustrative examples-pupil practice routine). More importantly, they cited the challenge of making the question as difficult as possible as a motivating factor. With the self-constructed test, the pupils enjoyed making up challenging problems and questions for their classmates to solve. Indirectly, where the pupils were able to formulate the questions and problems in mathematically correct ways, they were revising their work spontaneously and setting their own learning target as well as challenge their own learning.

In short, we would say that the self-assessment strategies integrated into the Mathematics Classroom will help the pupils to be aware of their own learning and their teachers to know their pupils better. Although the research result on students' attitudes toward mathematics and the learning of mathematics is to a degree inconclusive, there is a clear sign that their attitudes are not worse than the comparison classes. As for the attitude towards the new strategy students and teachers generally welcome the ideas and are showing some significant positive change in attitude. Though the performance in new strategy do not show significant difference, the interview shows students and teachers believe in the help of the new strategy in a long run.

4.5 Implications and Recommendations

The findings of this sub-study recommend the use of self-assessment to be a viable assessment and instructional strategy in primary schools. Both teachers and pupils were rather uncomfortable in using the self-assessment tasks initially but "accepted as routine" with time. Self-assessment provides information for corrective feedback that is non-threatening because the self-assessment tasks are not marked and grades are not assigned. They are therefore most useful if it was immediate, frequent and communicated in non-judgmental ways.

The MAP team believes that as more and more teachers become knowledgeable about self-assessment and how they can begin using it, given time (as any changes will need time to see the results) self-assessment will move from being ad hoc activities to classroom practice. The impact of this section of the research project may be inconclusive because of unanticipated operational difficulties and because both pupils and teachers are new to the instruments. As long as both teachers and pupils can derive benefits from engaging in pupil self-assessment, there is high possibility that it can be implemented successfully in class. The task of establishing positive gains to pupils in mathematics performance from the use of pupil self-assessment would require further study.

Chapter 5 Results and Findings (III): Project Work (Primary)[6]

5.1 Introduction

This chapter focuses on the impact of project work or mini investigations on students' mathematics learning over about 18 months. As said earlier, it is part of the Mathematics Assessment Project (MAP) that was designed to investigate the effects of varied assessment strategies such as journal writing, project work/investigation, performance task and self assessment, on student learning as measured by the following:

- change in attitudes and beliefs towards mathematics learning,
- the ability to cope with open-ended tasks, and
- mathematics achievement.

In MAP, a project is defined as a task or a series of tasks for students to carry out, which generally includes some or all of the following processes: gathering data, observing, looking for references, identifying, measuring, analyzing, determining patterns and/or relationships, graphing or organizing data/information, communicating (In written or oral language).

An investigation is defined as the process of exploring and finding the solution, which is not readily available (and often open-ended), to a task. An investigation often includes observing, identifying, measuring, graphing, analyzing, and so on, therefore doing a project usually requires investigation. In this sense, we sometimes use project work and investigative work interchangeably.

5.2 Research Questions

The specific questions of this sub-study were as follows:

(a) What is the impact of investigation/project work on the attitude of students towards mathematics?

(b) What is the resulting attitude towards investigations/projects?

(c) Does investigation/project work impact performance?

(d) What suggestions can be made for implementing investigation/project work in primary classrooms?

5.3 Methods

5.3.1 Participating schools and classes

Two primary schools, a high-performing school and a neighborhood school, participated in the study. Three classes were selected from each school as shown in Table 5.1.

Table 5.1. Participating schools and classes in project-based assessment

	School A: High-performing	School B: Neighbourhood
Experimental /High	Average ability class (E1 & E2)	High ability class (E3)
Comparison / High	Comparison class for E1 & E2 (C1)	Comparison class for E3 (C2)
Experimental /Low	-----	Low ability class (E4)

The class sizes varied from year to year in School B as the students were streamed according to their performance, based particularly on their mathematics performance and language ability. In 2004, the students were in Primary 3 (P3) and in 2005, they were in Primary 4 (P4). In School A, the students from the two experimental classes have comparable ability with the students in the comparison class. This is not so in School B. Students in one of the experimental classes (Class E4) have very low ability in mathematics as well as poor language skills. Many students in this class seldom passed their mathematics achievement tests in school.

The teachers participating in the project also varied from year to year. All teachers participated in the study are teachers with teaching qualifications, some of them are degree holders while others are diploma holders.

In School A, the teachers, particularly the two teachers (Teachers T3 and T4) who taught Class E2, are relatively younger and less experienced than those in School B. The students had different teachers when they were promoted from Primary 3 to Primary 4. Table 5.2 presents the information of these teachers in School A.

Table 5.2. A profile of participating teachers in School A

	E1		E2		C1	
	P3	P4	P3	P4	P3	P4
Teacher	T1	T2	T3	T4	T5	T6
Qualification	PGDE	Dip Ed	degree	PGDE	PGDE	Dip ed
Gender	F	F	F	F	F	F
Age	36 - 40	31 - 35	25 - 30	21 - 25	36 - 40	31 - 35
Teaching experience (yrs)*	> 5	< 5	> 5	< 5	> 5	< 5

Note. *The school does not have the record of the exact years of teaching experience.

In School B, the experienced teachers taught the comparison classes and the younger teachers taught the experimental classes. In Primary 4, all classes were taught by teachers different from those in Primary 3. The teacher (Teacher T8) who taught E4 (a low ability class) in Primary 3 was reassigned to teach E3 (a high ability class) in Primary 4. It should be pointed out that the teacher (Teacher T10) who taught Class C2 in Primary 4 was also the Head of Department of Mathematics at the school.

Table 5.3 presents the information of these teachers in School B.

Table 5.3. A profile of participating teachers in School B

	E3		C2		E4	
	P3	P4	P3	P4	P3	P4
Teachers	T7	T8	T9	T10	T8	T11
Qualification	PGDE	PGDE	Cert.Ed	FPDE	PGDE	Dip.Ed.
Gender	F	F	M	F	F	F
Age	31-35	31-35	50+	41-45	31-35	21-25
Teaching experience (yrs)	5	4	37	20	4	4

5.3.2 Instruments and data collection

As for the whole MAP project, in addition to the mini projects/investigations (hereinafter referred to as investigations) given to the students in the experimental classes, a survey of attitudes towards mathematics and an open-ended test were also administered in all six classes (4 experimental classes and 2 comparison classes). The survey and the test were given before and after the intervention period. That is, when the students were in Primary 3 and later, when they were in Primary 4.

The following data were also collected:

- Achievement in mathematics as measured by the school semestral assessment in mathematics.

- Teacher perception on the use of investigations to assess student learning as revealed through teacher interviews.

- Student perception on mathematics learning as revealed through student interviews

Survey of attitudes towards mathematics
The survey was similar to all the components in the MAP project. It measures the student attitudes and beliefs towards mathematics in five scales:

- Scale G: General view toward mathematics and mathematics learning
- Scale A: Anxiety towards mathematics learning
- Scale P: Perception of performance in mathematics
- Scale B: Beliefs about the usefulness of mathematics

The same test was used in the pre-test and in the post-test. In addition, students in the four experimental classes were also asked to respond to 13 statements on project work/investigations.

Open-ended test
A project or a mathematics investigation is often open-ended in nature and the focus is on the thought processes rather than the products. It provides opportunities for students to solve problems in real-life context. The open-ended test allows the researchers to better understand the primary students' ability to deal with open-ended task and use their thinking in real-world situation to analyze data and make decisions. The test consists of four items:

- a problem with extraneous information;
- a problem with missing information;
- a problem involving data handling; and
- a problem with alternate answers.

The same tasks were presented in the pre-test and in the post-test, albeit in a different sequence. A copy of the tasks can be found in Appendix 5.1.

The mini projects/investigations (referred to as investigation)
Most of the investigations used in the study are open-ended tasks. A task is open because there is either more than one solution strategy or more than one answer. All the tasks were designed to match the curriculum. One task was designed for each unit in the mathematics syllabus for Primary 3 and Primary 4. The students were expected to complete most of the tasks in class, often within 2 periods (that is, 1 hour). It was noted that teachers used the tasks for different purposes at different time, for example, as a means to extend student learning or to review the relevant concepts and skills taught. For each task, teacher's notes, an evaluation rubric, and student reflection accompany the activity sheets. A check list of the investigation tasks and a sample of the resource materials for the investigation on the unit on Angle (Primary 4) can be found in Appendix 5.2 and Appendix 5.3 respectively.

Interviews
Using the same research methods as in the other parts of MAP, both the teachers and the students in the four experimental classes in this sub-study were interviewed at the end of the intervention. The students were interviewed in small groups using a set of structured interview questions. The interview was meant to provide additional information on the students' perceptions of the tasks they had done during the intervention period. Moreover, the information collected would triangulate the data collected in the survey.

Another set of structured interview questions was used in the teacher interview to collect information on the teachers' perceptions of investigations and their experience in carrying out the tasks during the instructional time, including the difficulties they faced.

As reported earlier, the outline of the structured interview questions for the teachers and the students can be found in Appendix 2.4 and Appendix 2.5 respectively.

5.4 Results

In this report, no attempt was made to compare the experimental low class (E4) and the comparison class (C2) in School B because the differences in the mathematics ability and language proficiency between these two classes are too wide for the comparison to be meaningful. Class E4 is the weakest class in the cohort. In this class, the teacher had to go over the prerequisites and the procedure of the investigation one step at a time. The tasks were often completed as a whole class discussion and were time consuming. It should be pointed out that some of the students in Class E4 could not even read the activity sheet.

5.4.1 Results from questionnaire surveys

There are five negative statements (Q7, 9, 10, 11 and 12) in the survey. The students' responses to these statements were recoded to reflect the rating towards the corresponding positive statements. In this section of the report on the survey, the filled bars in the figures represent the data collected when the students were in Primary 3 before they were given the first investigation (pre-survey) and the stripped bars represent the data collected at the end of the study(post survey) when the students were in Primary 4.

Scale G: Students' general views about mathematics and mathematics learning
There are five items in Scale G, that is, Q1, Q6, Q7, Q12 and Q14. Table 5.5 presents the distribution of the student responses to these items in the pre- and post survey for School A.

Table 5.5. Distribution of students' responses in Scale G in School A

School A		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q1	E1	5.2%	2.6%	92.1%	6.1%	15.2%	78.7%
	E2	5.6%	16.7%	77.7%	22.9%	22.9%	54.2%
	C1	5.6%	8.3%	86.1%	14.0%	16.7%	69.5%
Q6	E1	8.1%	2.7%	89.2%	12.1%	3.0%	84.9%
	E2	2.8%	8.3%	88.8%	20.0%	22.9%	57.2%
	C1	2.7%	8.1%	89.1%	22.2%	11.1%	66.7%
Q7	E1	23.7%	2.6%	73.7%	24.3%	18.2%	57.6%
	E2	27.8%	8.3%	63.9%	37.1%	8.6%	54.3%
	C1	35.1%	0.0%	64.8%	27.9%	19.4%	52.8%
Q12	E1	5.3%	5.3%	89.5%	18.1%	3.0%	78.7%
	E2	11.2%	8.3%	80.6%	22.9%	22.9%	54.3%
	C1	5.4%	10.8%	83.7%	14.0%	22.2%	63.9%

Q14	E1	8.1%	13.5%	78.3%	21.2%	21.2%	57.7%
	E2	22.3%	22.2%	55.6%	31.4%	31.4%	37.1%
	C1	16.2%	21.6%	62.2%	41.6%	25.0%	33.4%

Table 5.6 shows the distribution of the student responses to the items in Scale G in the pre- and post survey for School B.

Table 5.6. Distribution of students' responses in Scale G in School B

School B		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q1	E3	4.8%	12.2%	82.9%	10.3%	10.3%	79.6%
	C2	10.0%	5.0%	85.0%	17.1%	14.6%	68.3%
	E4	24.1%	27.6%	48.3%	28.0%	8.0%	64.0%
Q6	E3	2.5%	2.5%	95.0%	18.0%	10.3%	71.8%
	C2	10.0%	2.5%	87.5%	12.2%	9.8%	78.1%
	E4	16.1%	9.7%	74.1%	28.0%	8.0%	64.0%
Q7	E3	17.5%	22.5%	60.0%	12.9%	10.3%	76.9%
	C2	5.0%	12.5%	82.5%	22.0%	19.5%	58.6%
	E4	63.3%	6.7%	30.0%	40.0%	12.0%	48.0%
Q12	E3	9.7%	2.4%	87.9%	12.9%	2.6%	84.6%
	C2	7.5%	0.0%	92.5%	12.2%	12.2%	75.6%
	E4	26.7%	20.0%	53.3%	36.0%	0.0%	64.0%
Q14	E3	17.0%	9.8%	73.2%	30.8%	15.4%	53.8%
	C2	18.0%	2.6%	79.5%	26.9%	26.8%	46.4%
	E4	19.3%	9.7%	71.1%	24.0%	12.0%	64.0%

Figure 5.1 and Figure 5.2 show the mean rating scores for the pre- and post- survey ratings for the 4 experimental classes and the two comparison classes.

It is noted over the 18-month experimental period, all average and high ability students in both the experimental and comparison classes seemed to enjoy mathematics less (Q1), find mathematics less interesting (Q6), liked mathematics less (Q12) and were less likely to spend time in studying mathematics (Q14). Except for Class E3, the students also found mathematics is not hard for them (Q7).

The trend is not reflected for low ability students in Class E4 where more students reported in the post-survey than in the pre-survey that they enjoyed doing mathematics (Q1), mathematics is not hard for them (Q7), and they like mathematics (q12).

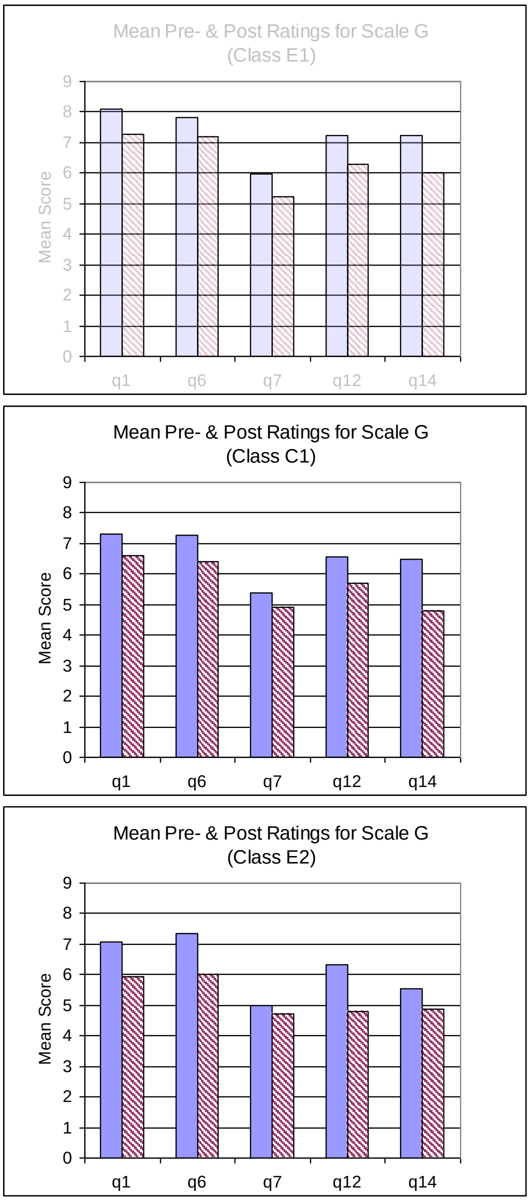
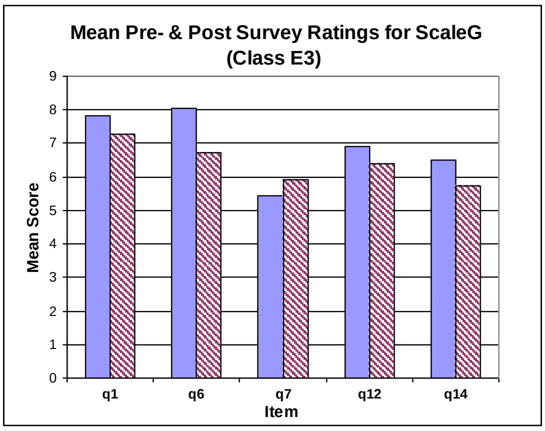
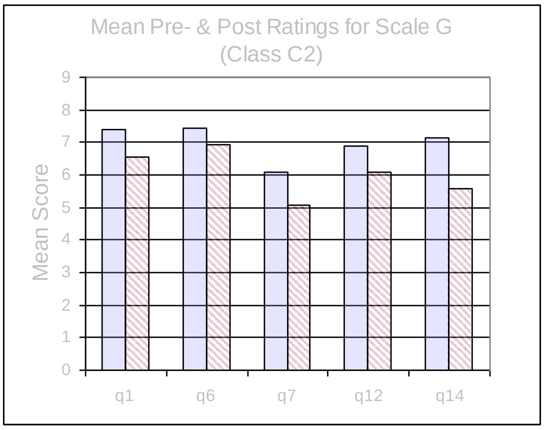


Figure 5.1. Mean Ratings in pre- and post-survey to items for Scale G (School A)





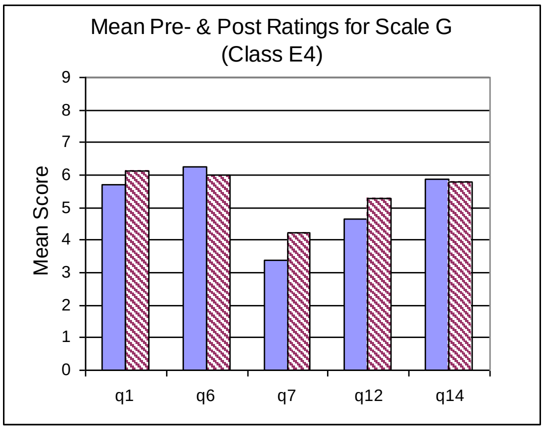


Figure 5.2. Mean Ratings in pre- and post-survey to items for Scale G (School B)

Figure 5.2 shows that compared to Class C2, Class E3 has relatively higher mean ratings for Q1 and Q6 in the pre survey. There is very little difference in the mean ratings between the two classes in item Q12. In the post survey, Class E3 has higher mean ratings than Class C2 in all items except Q6. This may be due to their exposure to the investigations given during the intervention period. These non routine tasks are not found in their textbooks or the activity books and often involve pair or group work. These bright students enjoyed the challenge and might find the textbook practices and worksheet exercises monotonous and not challenging. Hence, the routine school mathematics becomes less interesting to them. More and more students in this group are also finding mathematics not so difficult for them after all over time (Q7)

Figure 5.3 and Figure 5.4 present the difference in the mean ratings of items in Scale G between the experimental classes and the corresponding comparison class in the pre- and the post-survey. A positive value shows that the mean ratings for the experimental class is greater than that for the comparison class while a negative value indicates that the comparison class has a greater mean than the experimental class. The solid bars represent the differences found in the pre-survey while the stripped bars represent the differences found in the post survey.

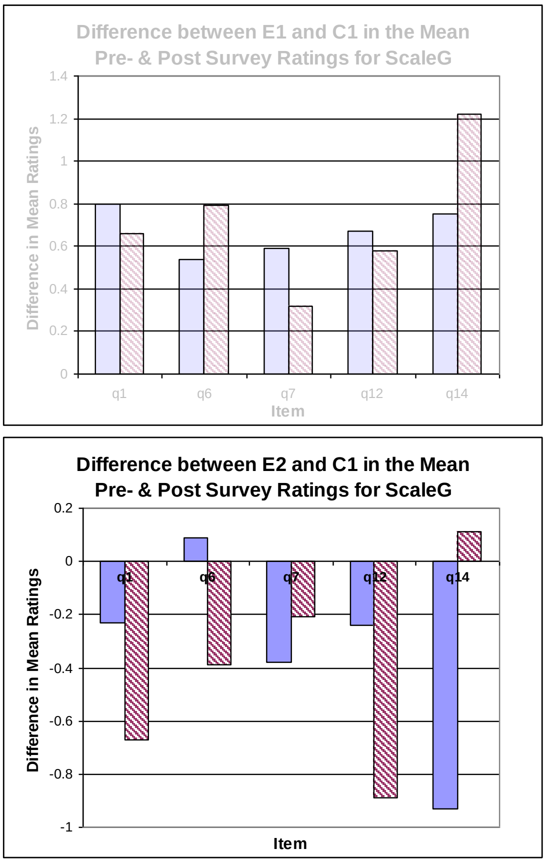


Figure 5.3. Difference between pre- and post-survey ratings on Scale G (School A)

Figure 5.3 shows that Class E1 consistently had more positive responses to all the five items in Scale G compared to the Class C1. However, the difference between in the mean ratings for Class E1 and Class C1 decreases in the post survey for items Q1, Q7 and Q12 but increases for items Q6 and Q14.

The pattern is different for Class E2. In the pre-survey, fewer students in Class E2 than in Class C1 responded favorably to all items except item Q6. The differences between the responses from the two classes increase in the post-survey except for items Q7 and Q14 where the students in Class E2 showed relatively more favorable responses in the post survey than the students in Class C1. Figure 5.4 also shows that Class E3 became more positive in the post-survey for all the items except items Q6.

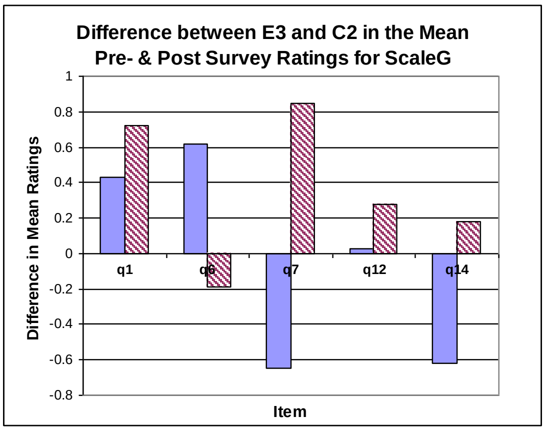


Figure 5.4. Difference between the experimental class and comparison class on Scale G (School B)

<u>Scale A: Students' anxiety towards mathematics learning</u>
There are three items in Scale A, Q2, Q9 and Q10. Table 5.7 and Table 5.8 show the distribution of the students' responses in these three items in the pre and post surveys for School A and School B respectively.

Table 5.7. Distribution of responses in Scale A for School A

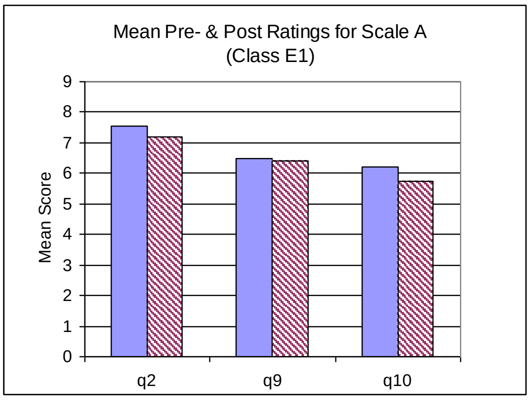
School A		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q2	E1	10.8%	5.4%	83.7%	9.0%	15.2%	75.8%
	E2	8.4%	2.8%	88.9%	17.1%	14.3%	68.5%
	C1	18.9%	2.7%	78.3%	22.3%	11.1%	66.7%
Q9	E1	15.7%	2.6%	81.6%	12.1%	12.1%	75.8%
	E2	13.9%	11.1%	75.0%	17.2%	22.9%	60.0%
	C1	21.6%	8.1%	70.2%	16.7%	27.8%	55.6%
Q10	E1	13.5%	5.4%	81.0%	30.2%	6.1%	63.6%
	E2	19.5%	16.7%	63.9%	22.9%	14.3%	63.0%
	C1	29.7%	5.4%	64.8%	8.4%	16.7%	74.9%

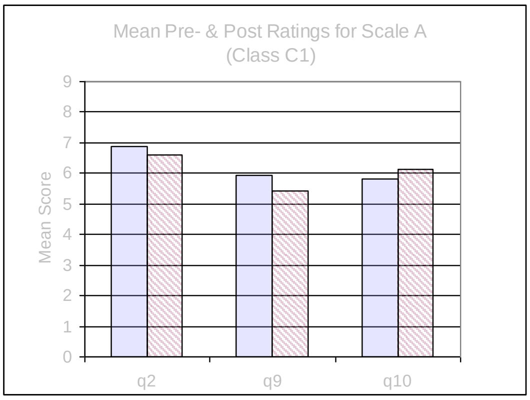
Table 5.8. Distribution of responses in Scale A for School B

School B		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q2	E3	9.7%	2.4%	87.9%	12.8%	7.7%	79.5%
	C2	5.0%	0%	95.0%	12.2%	24.4%	53.4%
	E4	6.7%	3.3%	90.0%	40.0%	8.0%	52.0%
Q9	E3	12.1%	2.4%	85.4%	18.0%	5.1%	76.9%
	C2	7.5%	7.5%	85.0%	12.1%	9.8%	78.1%
	E4	20.6%	17.2%	62.0%	32.0%	8.0%	60.0%
Q10	E3	17.1%	4.9%	78.1%	20.5%	17.9%	61.6%
	C2	22.5%	0%	77.5%	26.8%	17.1%	56.1%
	E4	31.0%	6.9%	62.0%	52.0%	0.0%	48.0%

From Table 5.7 and Table 5.8, we can see that generally, fewer students in the post-survey than in the pre-survey in both the experimental and the comparison classes reported that they were not afraid of doing mathematics (Q2) and felt good about mathematics (Q9). Besides one of the comparison classes (Class C1), the percentage of students who reported that it did not make them nervous to do mathematics also decreases in different amount for the other five classes in the post survey.

Figure 5.6 and Figure 5.6 show the mean rating scores for the pre- and post- survey ratings for the four experimental classes and the two comparison classes. The greatest difference between the mean ratings in the pre and post survey can be found in Item Q2 for Class E4 (difference = -1.35). Compared to other classes, there are relatively more students in this low ability class who are afraid of doing mathematics in the post survey than in pre-survey. The greatest improvement in the mean ratings can be found in Item Q9 for Class E4. These students seem to feel better about mathematics in the post survey than in the pre-survey (Q9).





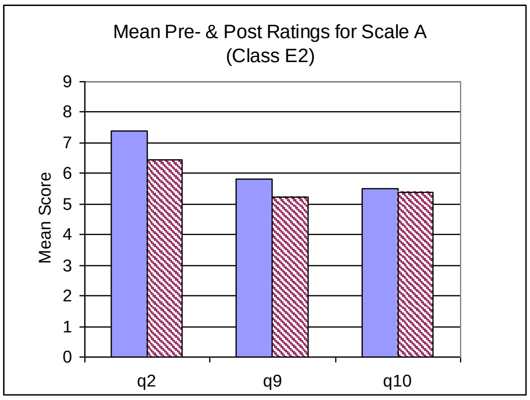
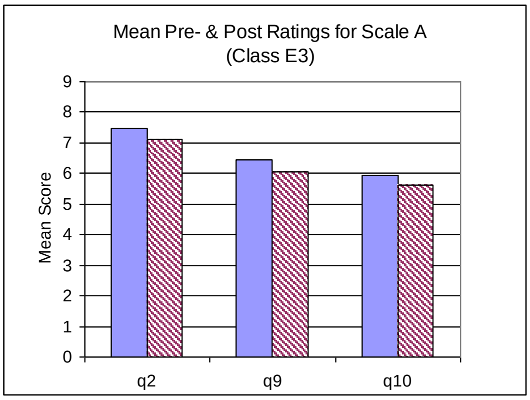
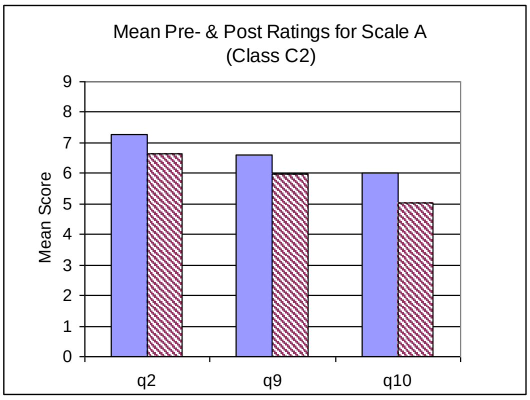


Figure 5.5. Mean ratings in pre- and post-survey to items for Scale A (School A)





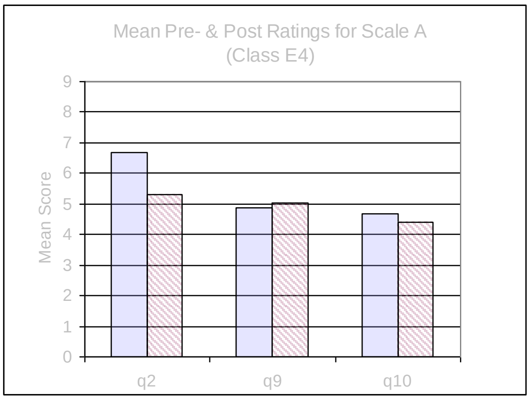
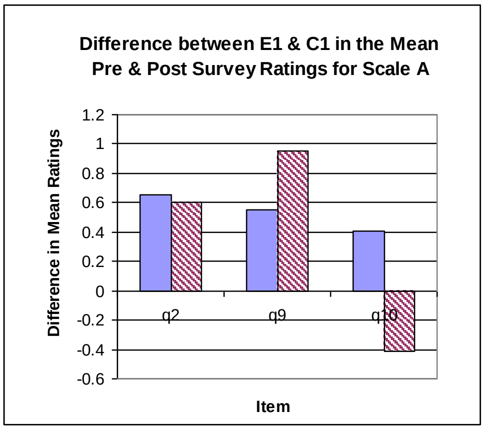


Figure 5.6. Mean ratings in pre- and post-survey to items for Scale A (School B)

The left part of Figure 5.7 shows that the difference in the mean rating between Class E1 and ClassC1 decreases in item Q2, and increases in Item Q9 from Primary 3 to Primary 4 with the means for Class E1 greater than that for Class C1. For item Q10, Class E1 has a higher mean rating than Class C1 at Primary 3 and a lower mean rating than Class C1 at Primary 4.

The right part of Figure 5.7 shows that the Class C1 has higher mean ratings than Class E2 in two of the three items at Primary 3 and in all three items at Primary 4. The difference in the mean ratings increases from Primary 3 to Primary 4.



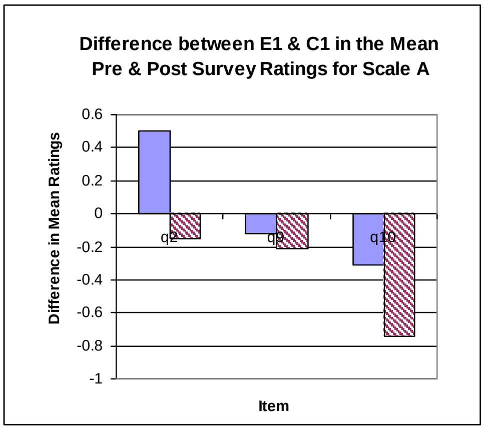


Figure 5.7. Difference between the experimental class and comparison class on Scale A (School A)

Fig. 5.8 presents the difference in the mean ratings between Class E3 and Class C2 for the three items in Scale A. Class E3 has a higher mean than Class C2 only in Item Q2 at Primary 3 but in all three items at Primary 4.

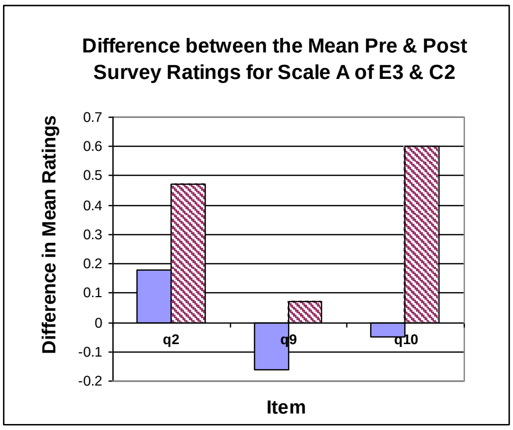


Figure 5.8. Difference between the experimental class and comparison class on Scale A (School B)

__Scale P: Students' perception of their mathematics performance__

There are four items in Scale P, Q3, Q4, Q11 and Q13. Table 5.9 and Table 5.10 show the distribution of the students' responses by percentages in these four items in the pre and post surveys for School A and School B respectively. Except for three instances, the percentages of favorable responses (Agree) to the four items generally decrease from the pre-survey to the post survey. In School A, a small increase in percentage is noticed for item Q4 (from 89.5% to 90.9%) for Class E1. In School B, there are small increases in percentage in Item Q13 (from 63.4% to 64.1%) and in item Q11 (from 61.0% to 65.9%) for Class E3, as well as a greater increase of about 20% points in Item Q4 for Class E4.

Table 5.9. Distribution of responses in Scale P for School A

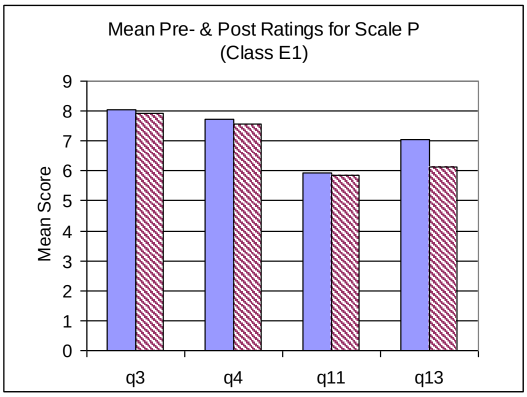
School A		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q3	E1	2.6%	7.9%	89.5%	0.0%	12.1%	87.9%
	E2	2.8%	8.3%	88.8%	14.3%	17.1%	68.6%
	C1	2.7%	2.7%	94.5%	5.6%	16.7%	77.9%
Q4	E1	2.6%	7.9%	89.5%	9.0%	0.0%	90.9%
	E2	5.6%	13.9%	80.5%	22.9%	17.1%	60.1%
	C1	8.1%	2.7%	89.1%	16.7%	16.7%	66.7%
Q11	E1	23.7%	2.6%	73.7%	21.2%	9.1%	69.7%
	E2	33.3%	8.3%	58.3%	48.7%	17.1%	34.3%
	C1	40.5%	8.1%	51.3%	44.4%	11.1%	44.5%
Q13	E1	18.4%	10.5%	71.1%	15.1%	21.2%	63.7%
	E2	36.1%	2.8%	61.1%	51.4%	17.1%	31.5%
	C1	29.7%	10.8%	59.4%	44.4%	8.3%	47.2%

Table 5.10. Distribution of responses in Scale P for School B

School B		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q3	E3	2.4%	4.9%	92.7%	5.2%	7.7%	87.1%
	C2	2.5%	5.0%	92.5%	12.2%	4.9%	82.9%
	E4	10.0%	0.0%	90.0%	12.0%	8.0%	80.0%
Q4	E3	2.5%	12.5%	85.0%	10.5%	10.5%	79.0%
	C2	10.0%	10.0%	80.0%	12.1%	9.8%	78.0%
	E4	3.2%	64.5%	32.3%	20.0%	28.0%	52.0%
Q11	E3	17.1%	22.0%	61.0%	26.3%	7.9%	65.9%
	C2	25.7%	7.7%	66.7%	34.2%	17.1%	48.8%
	E4	27.6%	6.9%	65.5%	64.0%	12.0%	24.0%
Q13	E3	19.4%	17.1%	63.4%	25.7%	10.3%	64.1%
	C2	22.5%	0.0%	77.5%	39.1%	17.1%	43.9%
	E4	38.8%	9.7%	51.6%	56.0%	12.0%	32.0%

Figure 5.9 and Figure 5.10 represents the mean ratings in all four items for all six classes in the two schools, respectively.

As we can find from Figure 5.9, in School A, the mean ratings to all the items in Scale P decrease from Primary 3 to Primary 4 for all the three classes (E1, E2 and C2). The greatest decrease can be found in Item Q3 (difference = -1.58) and Item Q13 (difference = -1.57) for Class E2. While in School B as shown in Figure 5.10, the largest increase is found in Item Q3 (difference = + 0.44) for Class E4 and the largest decrease can be found in Item Q11 (difference = - 1.2) also for Class E4.



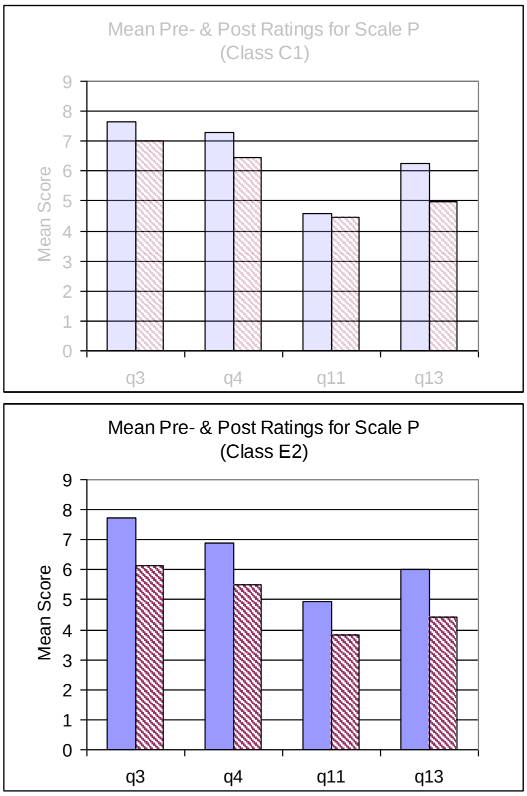
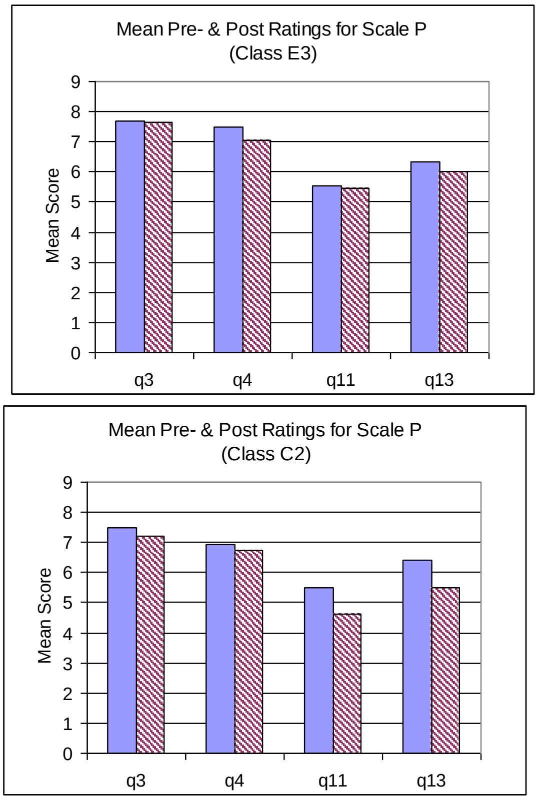


Figure 5.9. Mean ratings in pre- and post-survey to items for Scale P (School A)



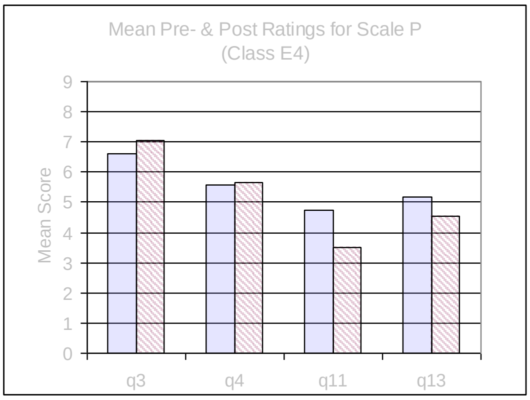


Figure 5.10. Mean ratings in pre- and post-survey to items for Scale P (School B)

Figure 5.11 depicts the mean ratings in all four items in the three classes in School A. As we can see that the difference in the mean ratings between Class E1 and Class C1 increases in all four items in Scale P from Primary 3 to Primary 4 with Class E1 having the higher mean ratings in all cases.

However, the pattern is not reflected in the comparison between the mean ratings for Class E2 and Class C1. At Primary 3, Class E2 has higher mean ratings than Class C1 for Items Q3 and 11. However, at Primary 4, Class C1 has higher mean ratings then Class E2 for all four items in Scale P. The difference in the mean ratings increases in all items from Primary 3 to Primary 4.

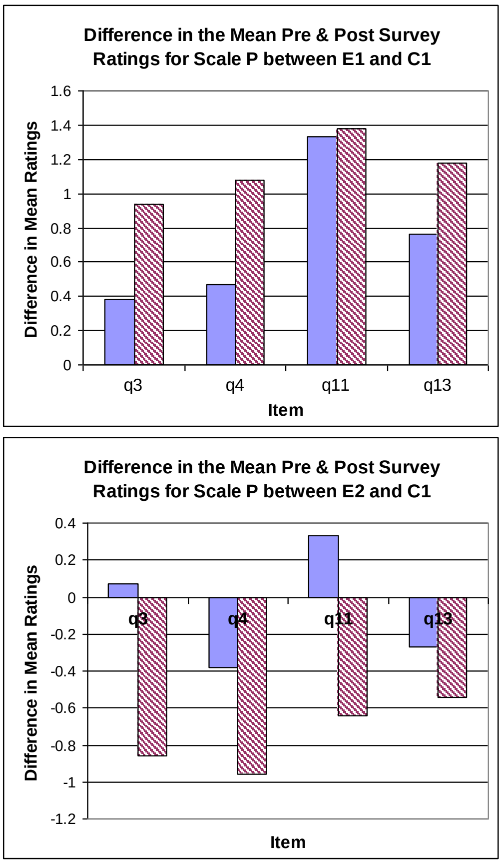


Figure 5.11. Difference between the experimental class and comparison class on Scale P (School A)

Figure 5.12 shows the mean ratings in all four items in Class E3 and C2 in School B. It can be found that the difference between the mean ratings for Class E3 and C2 increases from the pre-survey to the post survey for three of the four items as shown in Figure 12. The difference in the mean ratings for Item Q4 between Class E3 and Class C2 decreases from Primary 3 to Primary 4. Students in Class E3 seemed to have a relatively more positive perception of their mathematics performance than the students in Class C2. Their perception does not change as much as that of the students in Class C2.

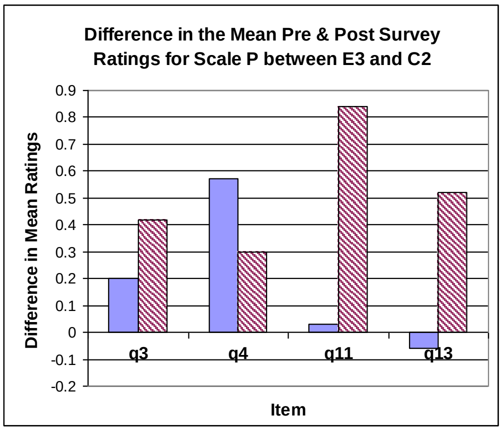


Figure 5.12. Difference between the experimental class and comparison class on Scale P (School B)

<u>*Scale B: Students' Belief about the usefulness of mathematics*</u>
There are two items in Scale B, Q5 (I think mathematics is useful to me.) and Q8 (Knowing mathematics will help me get a good job next time.) concerning the usefulness of mathematics.

Table 5.11 presents the distribution of the students' responses to these two items at Primary 3 (pre-survey) and at Primary 4 (post survey) for School A. Table 5.12 shows the corresponding data for School B.

Table 5.11. Distribution of responses in Scale B for School A

School A		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q5	E1	0.0%	0.0%	100%	3.0%	0.0%	97.0%
	E2	2.8%	5.6%	91.7%	0.0%	17.0%	83.0%
	C1	0.0%	0.0%	100%	2.8%	8.3%	88.9%
Q8	E1	5.2%	15.8%	78.9%	3.0%	9.1%	87.9%
	E2	17.1%	2.9%	80.1%	5.8%	20.0%	74.2%
	C1	8.1%	5.4%	86.4%	0.0%	22.2%	77.7%

Table 5.12. Distribution of responses in Scale B for School B

School B		Pre-survey			Post-survey		
		Disagree	Neutral	Agree	Disagree	Neutral	Agree
Q5	E3	2.4%	4.9%	92.6%	5.2%	10.5%	84.2%
	C2	7.5%	0.0%	92.5%	0.0%	7.3%	92.6%
	E4	17.2%	20.7%	62.0%	8.0%	4.0%	88.0%
Q8	E3	2.4%	4.9%	92.7%	10.3%	10.3%	79.4%
	C2	5.0%	5.0%	90.0%	7.2%	12.2%	80.5%
	E4	7.4%	11.1%	81.5%	0.0%	4.0%	96.0%

As we can find from Figure 5.11, in School A, over 90% of the students at Primary 3 agreed that mathematics is useful to them. However, the percentage for each of the three classes decreases at Primary 4. Only over 83% of the Primary 4 students agreed. At Primary 3, more than 78% of the Primary 3 students agreed that knowing mathematics would help them get a good job in future. Eighteen months later, fewer students in Class E2 and Class C1 agreed with the statement but more students in Class E1 agreed with the statement.

In School B (see Figure 5.12), the percentage of students who agreed with the statements in Item Q5 and Q8 decrease in Class E3 and Class C2, while the percentage of students in Class E4 who agreed with the statements increases across the grades. Apparently more and more of the low ability students realize the importance of learning mathematics.

Figure 5.13 shows the mean ratings in Scale B in the three classes of School A.

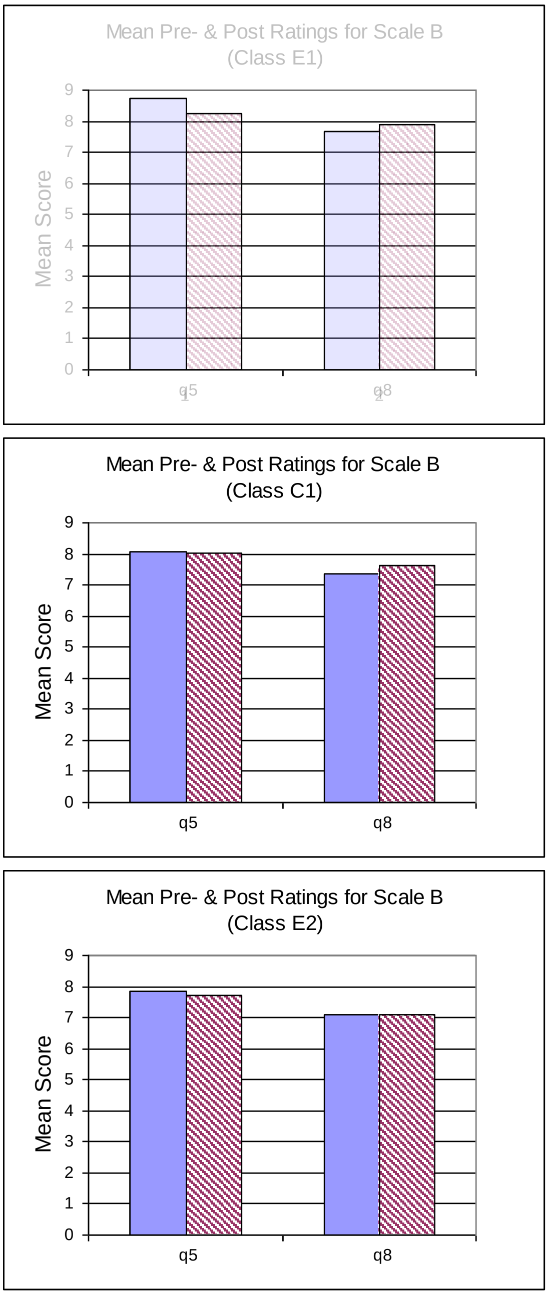


Figure 5.13. Mean ratings in pre- and post-survey to items for Scale B (School A)

As we can see from Figure 5.13, the changes in the differences of the mean ratings in all the classes seem relatively small.

Correspondingly, Figure 5.14 shows the mean ratings in Scale B in the three classes of School B.

It can be noted that the greatest decrease in the mean ratings can be found in Item Q5 for Class E3. The difference is 0.57 point. The greatest increase can be found in Item Q5 for Class E4. The difference is 1.46 points. The corresponding difference for Item Q8 is 1.45.

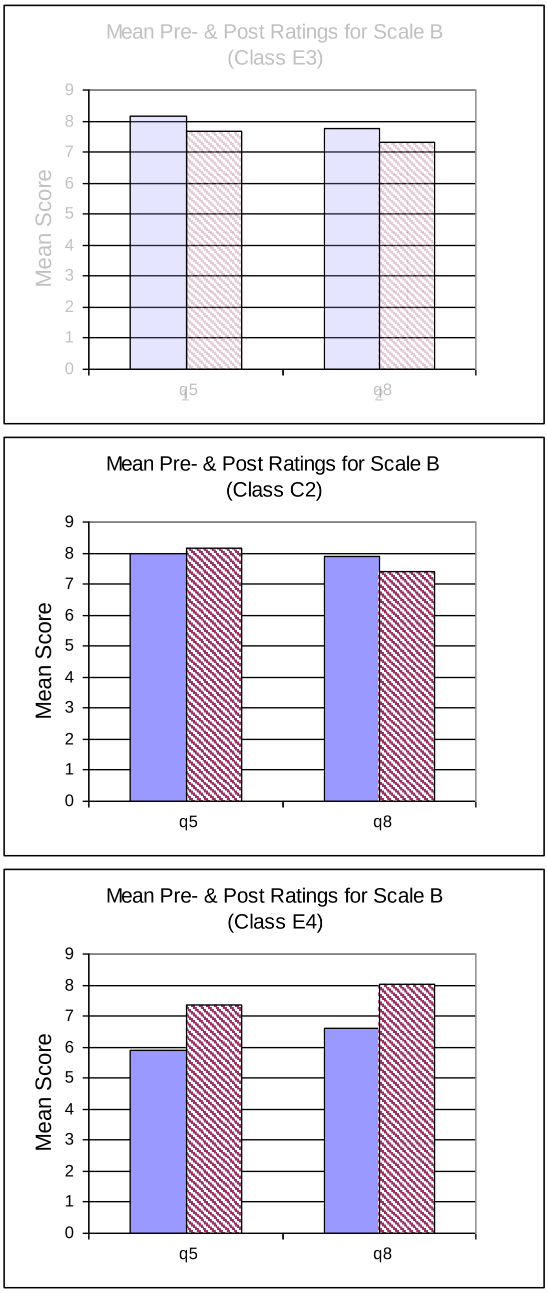
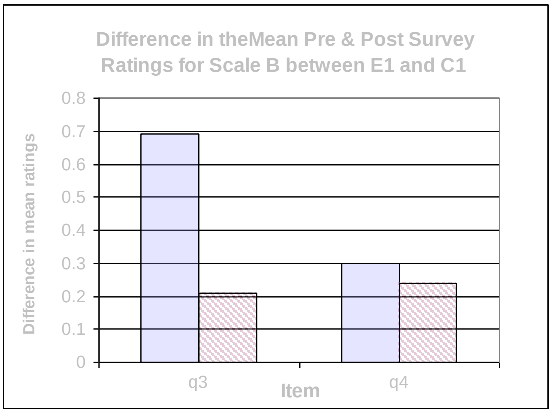


Figure 5.14. Mean ratings in pre- and post-survey to items for Scale B (School B)

Figure 5.15 presents the differences in the mean ratings between the experimental class and the comparison class in School A. It shows that Class E1 has greater mean ratings than the comparison class at Primary 3 and the differences diminish at Primary 4, while Class C1 has greater mean ratings than Class E2 at Primary 3 and the differences increase at Primary 4.



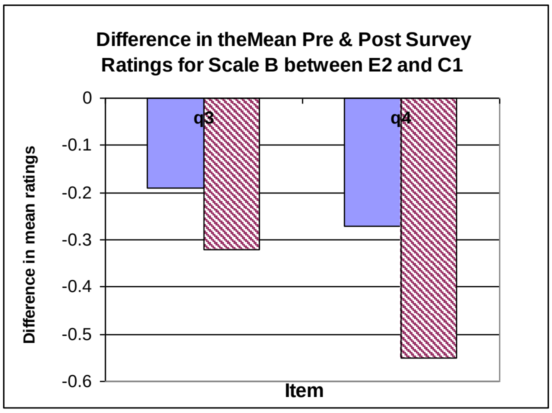


Figure 5.15. Difference between the experimental class and comparison class on Scale B (School A)

Figure 5.16 displays the differences in the mean ratings between the experimental class E3 and the corresponding comparison class C2 in School B. It can be noted that the Class E3 has a greater mean rating for Item Q5 than Class C2 at Primary 3. This difference is extended at Primary 4 with Class C2 having a higher mean ratings than Class E3. Class C2 has a greater mean rating for Item Q8 at Primary 4 but the difference diminishes at Primary 4.

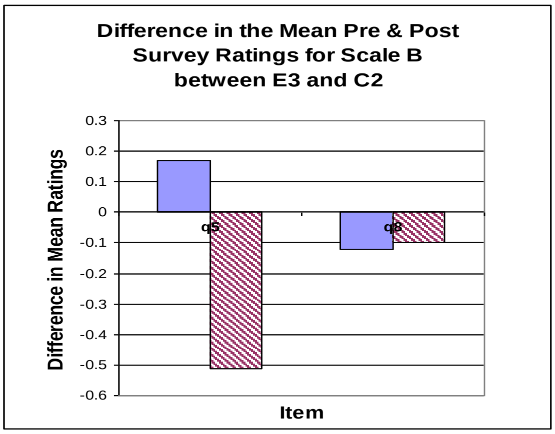


Figure 5.16. Difference between the experimental class and comparison class on Scale B (School B)

Students' attitudes towards project work

Table 5.13 shows the responses of students to the items concerning their attitudes towards project work. The negative statements in Q16, Q20, Q22 and Q27 have been recoded. (i.e., "disagree" to the negative statement is recoded as "agree" to the corresponding positive statement).

Table 5.13. Students' responses to items on attitudes towards project work

Item		School A		School B	
		E1	E2	E3	E4
Q15	Agree	75.8%	51.4%	48.7%	80.0%
	Neutral	12.1%	25.7%	20.5%	0%
	Disagree	12.1%	22.9%	30.8%	20%
Q16*	Agree	66.7%	60.0%	53.9%	48.0%
	Neutral	15.2%	14.3%	23.1%	12.0%
	Disagree	18.2%	25.7%	23.2%	40.0%
Q17	Agree	84.8%	74.2%	82.0%	84.0%
	Neutral	12.1%	11.4%	5.1%	4.0%
	Disagree	3.0%	14.4%	12.9%	12.0%
Q18	Agree	69.8%	39.9%	56.3%	76.0%
	Neutral	18.2%	34.3%	20.5%	8.0%
	Disagree	12.1%	25.8%	23.1%	16.0%
Q19	Agree	78.8%	65.8%	64.1%	80.0%
	Neutral	6.1%	31.4%	15.4%	16.0%
	Disagree	15.1%	2.9%	20.6%	4.0%
Q20*	Agree	63.6%	57.1%	74.3%	68.0%
	Neutral	15.2%	25.7%	7.7%	8.0%
	Disagree	21.1%	17.2%	17.9%	24.0%
Q21	Agree	78.9%	74.3%	71.8%	80.0%
	Neutral	18.2%	20.0%	15.4%	4.0%
	Disagree	3.0%	5.8%	12.9%	16.0%
Q22*	Agree	36.5%	34.3%	38.5%	20.0%
	Neutral	18.2%	22.9%	15.4%	40.0%
	Disagree	45.5%	43.0%	46.2%	76.0%
Q23	Agree	66.7%	37.1%	48.7%	45.8%
	Neutral	21.2%	34.3%	25.6%	29.2%
	Disagree	12.1%	28.6%	25.6%	25.0%
Q24	Agree	51.6%	48.6%	51.3%	52.0%
	Neutral	24.2%	37.1%	25.6%	16.0%
	Disagree	24.2%	14.3%	23.1%	32.0%
Q25	Agree	54.6%	48.6%	56.4%	84.0%
	Neutral	24.2%	22.9%	15.4%	8.0%
	Disagree	21.2%	28.6%	28.2%	8.0%
Q26	Agree	78.9%	77.1%	71.8%	84.0%
	Neutral	9.1%	14.3%	10.3%	4.0%
	Disagree	12.1%	8.6%	17.9%	12.0%
Q27*	Agree	69.7%	74.3%	64.1%	76.0%

	Neutral	18.2%	11.4%	12.8%	4.0%
	Disagree	12.1%	14.4%	23.1%	20.0%

As we can see, even though the mathematical ability of the students in the four experimental classes is different, the majority of the students in each class seemed to be aware of the benefits of doing mathematics project work, as highlighted below:

- Doing mathematics project work help me to learn mathematics (Q17, more than 74.2%)
- Doing mathematics project work help me to be more organized when I am doing mathematics (Q19, more than 64.1%),
- Doing mathematics project work make me more aware of the importance of learning (Q21, more than 71.8%)
- Working on mathematics project work makes me learn mathematics better. (Q26, more than 77.1%)
- Working on mathematics project work is not a waste of time. (Q27, more than 64.1%)

However, there are differences between the classes in their responses to the following statements on the liking of project works. Not all classes reported a favorable response (>50%) to these statements:

- Only 48.7% of students in Class E3 *like to do mathematics project work* (Q15),
- Only 39.9% of students in Class E2 *enjoy collecting my own data when doing mathematics project work* (Q18),
- Only 48.6% of students in Class E2 *would like to work on more mathematics project work for their mathematics lesson* (Q25).

Most of the students in Class E1 and Class E4 seemed to like project work. Moreover, generally, students were not very confident in doing mathematics project work, especially students in the low ability class (Class E4).

- Only 48% of the students in Class E4 claimed that doing mathematics work is not difficult to them (Q16),
- Few students (20% - 38.5%) agreed that they do not need more guidance when they are doing mathematics project work (Q22),
- Less than 50% of the students in three of the classes (37.1% for Class E2, 48.7% for Class E3 and 45.8% for class E4) thought that they do well in mathematics project work. (Q23).

There is not much difference in the percentages of students who think that *doing mathematics project work takes them more time than doing other mathematics questions usually done in class* (Q24). The percentages range from 48.6% in class E2 to 52% in Class E4.

The mean ratings were computed for each class. The maximum value possible is 9. Figures 5.17 to 5.19 present the means in three clusters of items, respectively.

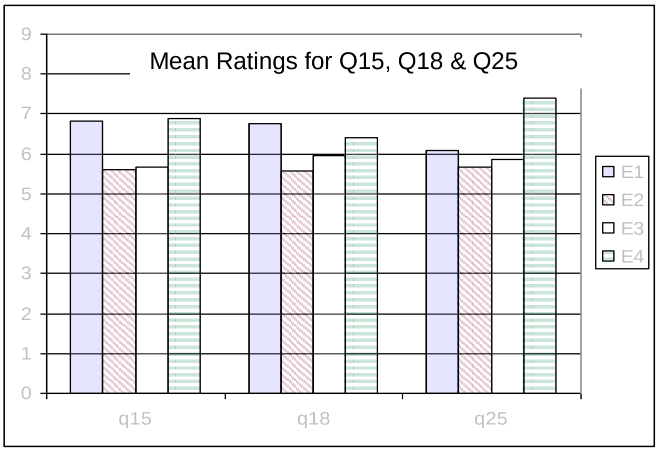


Figure 5.17. Mean ratings for items on mathematics project work for the four experimental classes (Items Indicating Liking Mathematics Project Work)

As shown in Figure 5.17, the first cluster of items gives a measure of how the students like mathematics project work. Apparently, the best class (E1) in the high performing school and the weakest class (E4) in the neighbourhood school like and enjoy doing project work than the other two experimental classes. Class E4 also has the highest mean in Item q25.

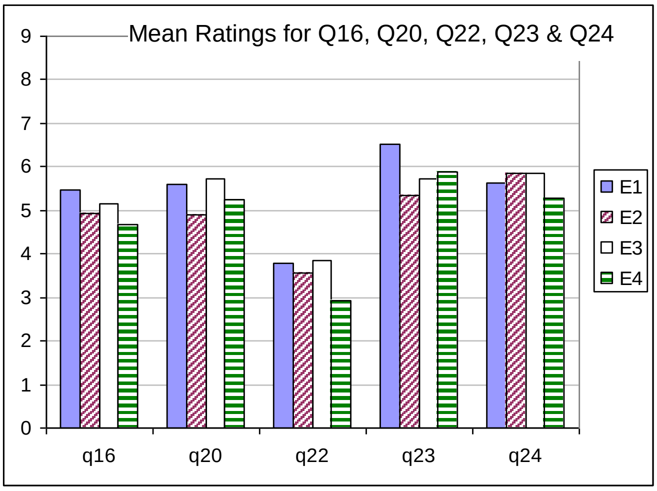


Figure 5.18. Mean ratings for items on mathematics project work for the four experimental classes (Items dealing with Difficulties of Project Work)

The second cluster of items deals with the difficulties of project work, shown in Figure 5.18. As expected, the ability of the students largely determines how they perceive mathematics project work. Class E4 has the lowest mean for three of the five items. Compare to the average students in ClassE1, E2 and E3, students with low ability (Class E4) tend to view project work as difficult (Q16). They feel that they need more guidance when doing project work (Q22), and need to spend more time to complete project work than other mathematics tasks (Q24). The class does not have the lowest means for Q20. It is probably because the class teacher, who is aware of the ability of her students, always provides a lot of scaffolding to help her students complete the task.

Consequently, the students also thought that they *do well in mathematics project work (Q23).* Even though Class E4 is a class with low ability students, their mean rating for Q23 is the second highest among the four classes.

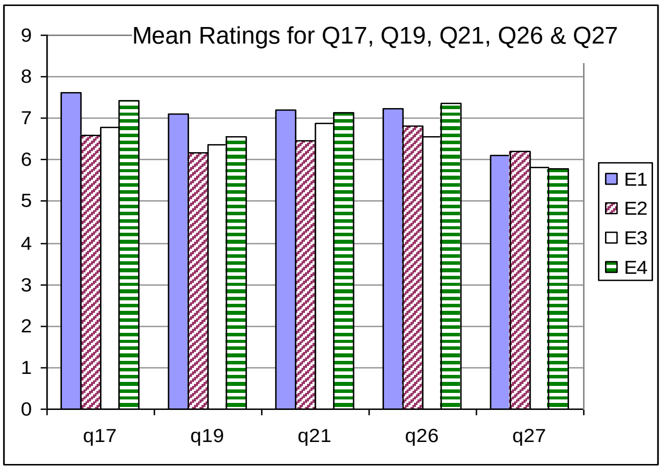


Figure 5.19. Mean ratings for items on mathematics project work for the four experimental classes (Items Dealing with Usefulness of Project Work)

Figure 5.19 depicts the corresponding results about the third cluster of items, which are about students' views on the usefulness of doing project work. Interestingly, Class E4 has the highest mean in Item Q26 among the four classes. They think that project work makes them learn mathematics better. Overall, one can see from the results that the students are very positive about the usefulness of the project work.

5.4.2 Results from pre- and post-tests

Participating students

Table 5.14 shows the number of students who completed both the pre-test and the post-tests. The students were streamed at the end of Primary 3. Consequently, only students remained in the same Experimental/High, Comparison/High and Experimental/Low class at Primary 4 were included.

Table 5.14. The number of students taking both pre- and post-tests

	School A			School B		
Class	E1	E2	C1	E3	E4	C2
No. of students	34	34	34	32	15	23

Test Items

Table 5.15 shows the test items and their characteristics. The items in the post-test are identical to those in the pre-test but placed in a different sequence. In the post test, item 2 was the item on data collection and item 3 was the item with insufficient information.

Table 5.15. Pre- and post-test items and their characteristics

	Items	Characteristics
1	Cik Fatimah bought 3 kg of sugar and 5 kg of flour. She bought 4 times as much rice as flour. How much rice did she buy?	This problem can be solved. However, it has extraneous information. Students need to distinguish the relevant information from the irrelevant information in order to solve the problem.
2	The refrigerator costs 5 times as much as the rice cooker. What is the total cost of the refrigerator and the rice cooker?	This problem has insufficient information. It cannot be solved as the cost of the rice cooker / refrigerator is not given.
3	How would you find out which is the favourite drinks of the Primary 3 students in your school? Write down the steps you would take to find out.	There is no numerical answer for this question. Students are expected to describe actions taken to collect data, organize data, represent the data and interpret the data.
4	Aini has only 50 ¢, 20¢ and 10¢ coins in her purse. She has 10 coins altogether. The total amount of the coins is $2. How many 10¢ coins does she have?	This is an open-ended problem. There is more than one solution. Students are expected to use any strategy to find the answers.

Analysis of test performance

Only the performance of students who completed both tests in the same class is analysed. The responses in bold in Tables 5.16, 5.17 and 5.18 are the percentage of correct responses.

1. Performance in Q1

Table 5.16 shows the distribution of the responses from the students in all six classes.

Table 5.16. Distribution of responses to Q1 in pre- and post-tests

School A	E1		E2		C1	
	Pre	Post	Pre	Post	Pre	Post
20 kg or 20 000 g or both	**91.2%**	**70.5%**	**85.2%**	**82.3%**	**82.3%**	**64.7%**
mention the mass of rice.			2.9%			
20 kg and an incorrect answer				2.9%	5.9%	8.8%
indicate that the problem can be solved. Give incorrect answer	8.8%	11.8%		8.8%	2.9%	
incorrect conversion		2.9%	5.9%			20.6%
use incorrect operation		5.9%	5.9%	2.9%	8.8%	2.9%
give nonsensical responses						2.9%
No response		8.8%		2.9%		
School B	**E3**		**C2**		**E4**	
	Pre	Post	Pre	Post	Pre	Post
20 kg or 20 000 g or both	**77.4%**	**93.8%**	**65.2%**	**69.5%**	**-**	**13.3%**
mention the mass of rice.	6.5%	6.3%	13.0%	17.4%	-	73.3%
20 kg and an incorrect answer	-	-	-	-	-	-
indicate that the problem can be solved. Give incorrect answer	3.2%	-	4.3%	-	7.1%	-
incorrect conversion	12.9%	-	-	-	-	-
use incorrect operation	-	-	13.0%	-	14.3%	-
give nonsensical responses	-	-	4.3%	-	35.7%	-
No response	-	-	-	13.0%	42.9%	13.3%

It should be pointed out that Q1 is an item that contains extraneous information. It is an item unfamiliar to the students as items with extraneous information are seldom found in the textbooks or activity books.

The results in Table 5.16 suggest that the performance of students in School A deteriorates in the post test for both the two experimental classes (E1 and E2) and the comparison class (C1). More students in Class E1 and Class E2 gave partially correct or incomplete answers in the post test while more students in Class C1 did so in the pre-test. Some students in Class E1 and Class C1 were not sure about conversion of units in the post-test. In fact more than 20% of the students in Class C1 committed this type of error.

On the other hand, performance of students in all three classes in School B improves in the post test. The greatest improvement occurred in the two experimental classes, from 77.4% to 93.8% in Class E3 and from 0% to 13.3% in Class E4. Apparently, relatively more students in the experimental classes than in the comparison class have developed the ability to identify the relevant information in problem solving. They realized that, unlike the routine problems found in the textbooks and activity books, not all information given in the problem was necessary to solve the problem. Moreover, there is also less variety of incorrect responses in the post-test for all the three classes. Students in Class E4 also made a greater effort in answering Q1 as the percentage of no responses drops from 42.9% in the pre-test to 13.3% in the post test

Comparing the performance of the students in the two schools, it is apparent that more students in School B have difficulty comprehending the question statements compared to the students in School A. Question 1 asks for the amount of rice bought. None of the students in School B gave the response “mass of rice” while 6.3%, 73.3% and 17.4% of the students did so in Class E3, Class E4 and Class C2 respectively.

2. Performance in Q2

Q2 cannot be solved as it has missing information. This is a non-routine task for the students. It requires students to identify the missing information. There is more than one possible correct response. The required missing information can be either the cost of refrigerator or the cost of the rice cooker. The structure of the given sentence *the refrigerator costs 5 times as much as the rice cooker* led most of the students to identify the cost of the rice cooker as the missing information, particularly in the post-test (see Table 5.17). Generally, student performance in this question improved in the post test and there is less variety in the responses found in the post-test than in the pre-test.

Table 5.17. Distribution of responses to Q21 in pre- and post-tests

School A	E1		E2		C1	
	Pre	Post	Pre	Post	Pre	Post
No. cost of refrigerator	**11.8%**	**5.9%**	**11.8%**	**11.8%**	**11.8%**	**14.7%**
No. cost of rice cooker	**70.6%**	**76.5%**	**76.5%**	**70.6%**	**70.6%**	**67.6%**
No. cost of refrigerator or cost of rice cooker	**8.8%**	**8.8%**	**2.9%**	**14.7%**	**2.9%**	**14.7%**
No. *cost of refrigerator & cost of rice cooker*					2.9%	
No. there is no number / there is not enough information	2.9%	2.9%			2.9%	
Yes. Answer given is often a multiple of 5	5.9%	5.9%	2.9%			2.9%
give nonsensical responses			2.9%	2.9%	8.8%	
No response			2.9%			
School B	E3		C2		E4	
	Pre	Post	Pre	Post	Pre	Post
No. cost of refrigerator	**3.2%**	**-**	**8.7%**	**4.3%**	**-**	**26.7%**
No. cost of rice cooker	**80.6%**	**96.9%**	**73.9%**	**95.7%**	**-**	**33.3%**
No. cost of refrigerator or cost of rice cooker	**6.5%**	**-**	**4.3%**	**-**	**-**	**-**
No. *cost of refrigerator & cost of rice cooker*	3.2%	-	4.3%	-	-	-
No. there is no number / there is not enough information	6.5%	-	-	-	7.1%	-
Yes. Answer given is often a multiple of 5	-	-	4.3%	-	7.1%	-
give nonsensical responses	--	-	4.3%	-	42.9%	-
No response	-	3.1%	-	-	42.9%	40.0%

As can be seen in Table 5.17, in School A, there is no difference in the percentage of correct responses in the pretest and the posttest for Class E1. There is an increase in the percentage of correct responses in both Class E2 and Class C1.

In School B, students in all the three classes performed better in the post test than in the pretest, particularly Class E4 with students of low ability. No students in Class E4 gave a correct response in the pretest but 60% gave a correct response in the post test. However, there is still a relatively large percentage (40%) of students in Class E4 who did not respond to this question.

Apparently by Primary 4, most of the students with average or high ability are aware that not all problems they come across have an answer. They need to read and understand the problem situation, identify the “known” and the “unknown” and whether the “unknown” can be computed from the “known” or is there some missing information. This ability would help them deal with the open nature of projects and investigations.

3. Performance in Q3

There is no numerical answer to this item. Students are required to consider a simple statistical investigative task and identify the steps required to answer the given question on data handling. Table 5.18 shows the different types of responses given by the students. The responses are classified according to the four stages in data handling: Data collection, Data organization, Data representation and Data interpretation. Responses are coded at the “highest stage” mentioned by the students. For example, response that mentions data collection and data organization is coded as data organization.

Generally, the students in both the experimental classes and the comparison classes performed better in the post test. The pattern of response is as expected because at the time of pretest, students were aware of picture graphs only and had limited experience on data handling. By the time they were in Primary 4, students were taught the bar graphs, particularly the interpretation of the bar graphs, and most of them were exposed to some activities on data collection during classroom instruction.

The noticeable improvement can be found in Class E4. The students in the class gave either nonsensical response or no response in the pretest. However, about 40% of the students were able to mention activities related to one or more of the stages of data handling in the post test. 60% of the students still gave no response in the post test. This is the highest percentage of no response among the six classes.

Table 5.18. Distribution of responses to Q3 in pre- and post-tests

School A	E1		E2		C1	
	Pre	Post	Pre	Post	Pre	Post
the step on data collection	38.2%	2.9%	35.3%	32.4%	64.7%	38.2%
the step on data organization/ tabulation/counting	23.5%	35.3%	17.6%	11.8%	11.8%	8.8%
data representation	2.9%	8.8%	2.9%	29.4%	8.8%	23.5%
interpretation		2.9%	2.9%	11.8%		2.9%
mention 1 or 2 of the above but in incorrect sequence	11.8%	29.4%	14.7%	11.8%		23.5%
nonsensical responses	11.8%		17.6%	2.9%	2.9%	
no response			8.8%			2.9%
School B	**E3**		**C2**		**E4**	
	Pre	Post	Pre	Post	Pre	Post
the step on data collection	61.3%	6.3%	8.7%	21.7%	-	-
the step on data organization/ tabulation/counting	12.9%	12.5%	73.9%	21.7%	-	26.7%
data representation	-	34.4%	4.3%	4.3%	-	-
interpretation	6.5%	28.1%	-	13.0%	-	6.7%
mention 1 or 2 of the above but in incorrect sequence	6.5%	15.6%	4.3%	21.7%	-	6.7%
nonsensical responses	9.7%	-	4.3%	-	50%	-
no response	3.2%	3.1%	26.1%	17.4%	50%	60%

4. Performance in Q4

Q4 is an open-ended item, a non-routine task for the students. There is more than one correct answer to this item. Students are allowed to use any strategy to solve this problem. However, they are expected to compute accurately. Table 5.19 shows the distribution of the student responses in the pre-test and the post-test. Except for Class E1, more students in the other five classes gave either one or two correct answers in the post-test than in the pre-test.

Table 5.19. Distribution of responses to Q4 in pre- and post-tests

School A	E1		E2		C1	
	Pre	Post	Pre	Post	Pre	Post
one correct answer, i.e., 3 or 6	**20.6%**	**17.7%**	**2.9%**	**32.4%**	**8.8%**	**11.8%**
2 correct answers. 3 and 6	**5.9%**	**2.9%**		**5.9%**		**8.8%**
3 or 6 and an incorrect answer	14.7%	14.7%	8.8%	8.8%	14.7%	26.5%
wrong answer including 13	44.1%	44.1%	50.0%	8.8%	44.1%	26.5%
how many 50¢ / 20¢ coins?	8.8%	20.6%	14.7%	38.2%	8.8%	23.5%
20 10¢ coins	2.9%		17.6%	2.9%	14.7%	2.9%
80¢	2.9%			2.9%		
nonsensical responses			2.9%		2.9%	
no response			2.9%		5.9%	
School B	**E3**		**C2**		**E4**	
	Pre	Post	Pre	Post	Pre	Post
one correct answer. i.e., 3 or 6	**35.5%**	**18.8%**	**8.7%**	**69.5%**	**-**	**33.3%**
2 correct answers. 3 and 6	**3.2%**	**75.0%**	**-**	**26.1%**	**-**	**-**
3 or 6 and an incorrect answer	12.9%	-	4.3%	-	-	-
wrong answer including 13	16.1%	-	26.1%	-	7.1%	-
how many 50¢ / 20¢ coins?	22.6%	-	17.4%	-	-	-
20 10¢ coins	3.2%	-	13.0%	-	14.3%	-
80¢	-	-	-	-	14.3%	-
nonsensical responses	6.5%	-	30.4%	-	14.3%	-
no response	-	6.3%	-	4.3%	50.0%	66.7%

The results in Table 5.19 show that, in School A, the percentage of correct responses decreases from 26.5% in the pre-test to 20.6% in the post test for Class E1. The corresponding percentage increases from 2.9% to 37.9% in Class E2 and from 8.8% to 10.6% in Class C1. None of the students in Class E2 and Class C1 gave both answers in the pre-test but 5.9% and 8.8% of the students in Class E2 and Class C1 respectively are able to do so in the post-test.

In School B, 35.5% 8.7% and 0% of the students in Class E3, Class C2 and Class E4 respectively seemed unaware of the alternate solution and gave only one correct answer to the question in the pre-test. In the post-test, 75%, 26.1% of the students in Class E3 and Class C2 were able to give both answers correctly but none of the students in Class E4 were able to do so. Different incorrect solutions can be found in the pre-test but not in the post test. Apparently, the students were either able to solve the problem or reluctant to attempt the question if they were not sure of the solution. Consequently, there was higher percentage of no response in the post test for all the three classes compared to the pre-test.

To summarize the results of students' performance on the tests as presented above, it appears that there is no conclusive finding about the impact of the intervention on student performance on the posttest. Nevertheless, we believe it is safe to say that overall there is neutral or positive influence on student performance. In particular, in School B where the investigations were given at regular intervals after each unit of study, the experimental class of low ability students performed better in all four items in the post test.

5.4.3 Results about the use of investigation tasks

There are more than one series of textbooks used in Singapore schools. School A and School B adopt different series and hence the schemes of work in these two schools are slightly different.

Nineteen investigative tasks were designed following the sequence found in Shaping Maths, the textbook series used in School A. However, not all tasks were carried out in both schools. School A began using the tasks in April 2004 and stopped administering the tasks in April in the following year. School B began using the tasks in May 2004. Class E1 managed to administer tasks by end of Term 2 (May 05) while Class E2 administered the last two tasks in Term 3. The four experimental schools did not administer all the tasks as planned due to various reasons including the different scheme of work, time constraints, preparation for examination, etc.

Table 5.20 lists the investigations completed by the students in the four experimental classes.

Table 5.20. Intervention tasks carried out in schools

		School A		School B	
Unit	Topic	E1	E2	E3	E4
P3A					
8	Multiplication & Division	May 04	May 04	April 04	April 04
P3B					
9	Mental Computation	July 04	July 04	May 04	May 04
10	Money	July 04	July 04	June 04	July 04
11	Length	July 04	Aug 04	July 04	July 04
12	Mass			Aug 05	Aug 04
13	Volume		Sept 04	Sept 04	Sept 04
14	Graph			Sept 04	Sept 04
15	Fraction	Sept 04			Nov 04
16	Time	Oct 04		Nov 04	
17	Geometry				
18	Perimeter & Area			Jan 05	Nov 04
P4A					
1	Whole Numbers	Mar 05	Feb 05	Jan 05	Jan 05
2	Multiplication & division	Mar 05		Feb 05	
3	Fractions	Mar 05	April 05		
4	Tables & Graphs	May 05	April 05	Feb 05	
5	Angles	May 05	April 05	April 05	April 05
6	Perpendicular & Parallel Lines	May 05		April 05	April 05
7	Geometrical Figures	May 05	Aug 05		
8	Area & perimeter	May 05	Aug 05		
	Total completed	14	11	14	12

Note. Some of these lessons are recorded on video.

In addition, some teachers did not administer the relevant task when they were teaching or when they just completed the unit. As they rushed to complete the syllabus and prepared their students for the end-of semester examination, they left the tasks for the students to carry out after the examinations. Generally, the “better” class in the schools, Class E1 in School A and Class E3 in School B, completed more tasks than the other experimental class in the school.

Sample of work
The following is a work sample for the unit on Angles in Primary 4. It is a sample from a pair of students in School B.

Objectives:
- To draw different triangles and 4-sided figures
- To measure angles using protractor

- To name and find angles that are less than 90^{o}, right angles, greater than 90^{o} and greater than 180^{o}
- To carry out an investigation and report on the investigation.

Task:
Students were given templates of 7-dot grid. There are required to draw different triangles on the grids, measure the interior angles using a protractor, and identify the angles greater or less than 90^{0}. They then extended their investigation to quadrilaterals (see Appendix 5.1).

Context:
Students were taught how to use the protractor to measure angles in the Unit on Angles. They had learnt that triangle is a three-sided shape and rectangles and squares are four-sided figures

Comment:
The work sample shows that the pair of students was able to draw different triangles and quadrilaterals. They were able to measure the angles using the protractor. In their report, they found that the sum of the angles in the triangle is 180^{0}, a property of the triangle found in the Primary 5 mathematics syllabus and the sum of angles in a quadrilateral is 360^{0}, a property not found in the primary mathematics syllabus. Not many students in both School A and School B note these properties, most of the students were preoccupied with the number of angles in a triangle and a quadrilateral. The students might have detected the relationships between the sum of the interior angles and the number of sides of the shape. They wondered what the sum of the angles is in a 5-sided figure and believed that it must be greater than 180^{0}.

In their reflection, the students pointed out that to complete the task properly, they need to read the correct scale in the protractor, understand and not just listen to the instructions.

This is a relatively well done sample of work. Most students in both schools were not aware of concave quadrilateral. The most common quadrilateral drawn is rectangles and rhombus, followed by trapezium and kite. Very few students drew the concave figure. A complete set of responses given by a pair of students is shown in Figures 5.20 to 5.22.

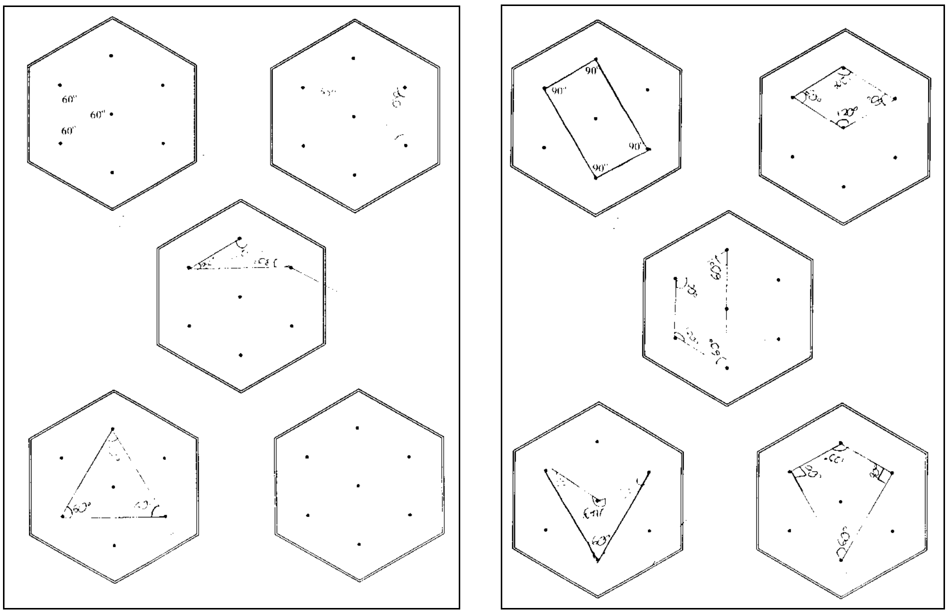


Figure 5.20. A sample of students' drawings


Report of our investigation

What we wanted to find out.

1. How many different triangles can be made on the seven-pointed grid? 3
2. How many different 4-sided figures can be made on the seven-point grid?
3. What types of angles are formed?

What we did:

First, we draw three-sided triangle. We measure the angles of the triangle. Next, we take down the degrees of the angles.

What we have found out

We found out that the three-sided triangle made up [illegible]. We also found out that all four-sided figures made up 360°.

What we wonder.

Five sided figure made up how many degrees. I think it is 180° more.

Figure 5.21. A sample of student report

Reflection: Angles

Did you …..	*Yes!*	*No!*	*I want to say that …*
➢ find measuring angles using a protractor difficult?		✓	I want to say that we must know which scale are we using.
➢ know exactly what you were supposed to do for the investigation?	✓		I want to say that know what are the things that we needed to use.
➢ listen to the instructions carefully?	✓		I want to say that we must understand but not just to listen.
➢ draw the figures neatly on page 3 and 4?	✓		
➢ write your report clearly on page 5?	✓		
➢ ask the teacher for help when you were not able to go on with the activity?	✓		

Figure 5.22. A sample of student reflection

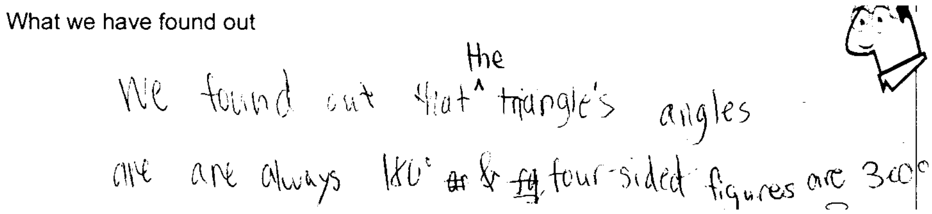
Generally, when asked to report what they had done, most students mentioned the actions of drawing and measuring. Many did not mention the use of protractor. Some students were observed to put aside the protractor after using it to measure a few angles. They seemed to be able to estimate the measures of the angles drawn on the 7-point grid after awhile. For example:

What we did:

1. We thinked what to draw
2. We tried it out.
3. We measured the angles.

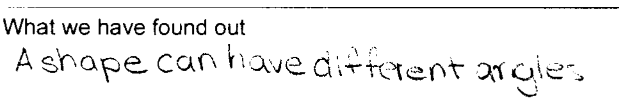
What we did: We drew triangles and 4-sided figures seven point grid and measured them.

When asked what they have found, there is a tendency for students to focus on the different angles measured or the different figures formed rather than the relationships between the measures of angles for the figures drawn. Very few students actually went further to find the sum of the angles in the shape as shown in Figure 5.20 and below.

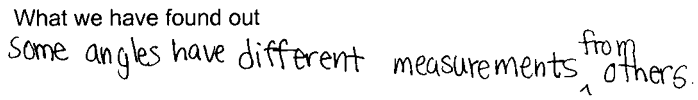

What we have found out

We found out that the triangle's angles are always 180° ~~or~~ & ~~fig~~ four-sided figures are 360°

The majority of the students did not know what to look for. They tend to mention what are obvious. For example:


What we have found out

A shape can have different angles


What we have found out

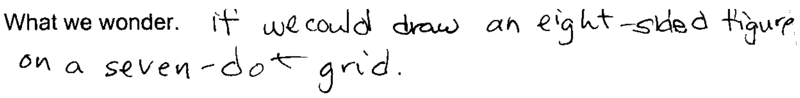
Some angles have different measurements from others.

Students' responses were varied when asked what they would like to know after completing the task. For example, some students wondered about the number of different triangles and quadrilateral that can be drawn on the grid.

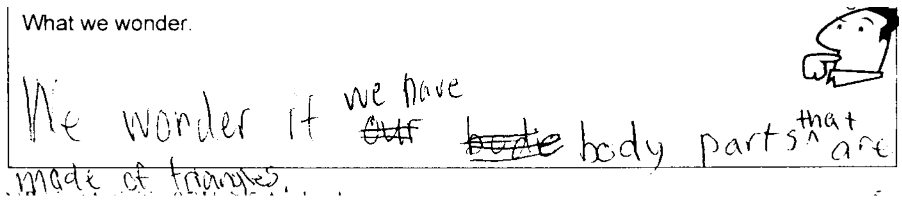

What we wonder.

How many more triangles we can draw.

Other wondered what happens if the number of sides of the figure is increased to eight.

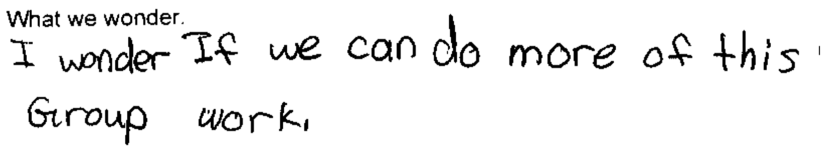

What we wonder. If we could draw an eight-sided figure on a seven-dot grid.

Other wondered where they can find triangles in real life situation. Below is such an example.


What we wonder.

We wonder if we have ~~our~~ ~~bode~~ body parts that are made of triangles.

There are also students who wonder about the task itself, why they are doing it or when can they have similar task involving group work again.


What we wonder.

I wonder If we can do more of this Group work.

This is not a difficult task for Primary 4 students. Students in Class E1, E2 and E3 did have little difficulty completing the task if they “know which scale of the protractor to use” and “understand but not just (to) listen” to the instruction given (see Figure 5.21). However students in Class E4 needed more scaffolding. With close guidance, most of them managed to complete the task. The following is the comment from the teacher for Class E4 after the lesson.

Overall, the lesson is within student’s understanding and ability and has provided opportunity for me to identify those who are still weak at measuring angles and reading the protractor efficiently.

5.4.4 Results from Interviews

Teacher interviews

The following are reports of how the four teachers teaching the Primary 4 classes viewed the mathematics investigation, the problems they faced when using the tasks and their recommendations for the use of investigative tasks in the primary mathematics classrooms.

1. View of the task

Teachers were aware that the investigative tasks were different from the routine tasks found in the textbooks and the activity books. The investigative tasks provided opportunities for group work and involve application of mathematics learnt. Many of the tasks are inter-disciplinary. The teachers appeared to like these features of the tasks. Usually, the students were enthusiastic.

> I like the fact that most of the tasks integrate other disciplines..., and there’s also elements of like games, and using concrete, real life examples... It (the task) creates more enthusiasm and to a large extent, helps the get them (students) to love mathematics. (Teacher E4)

> It (the investigation task) gives them... avenue for them to think logically... enhances their reasoning ability (Teacher E3)

The teachers noted the resource provided such as the accompany teachers’ notes.

> specific lesson plans that are given, like what are the objectives and how to accomplish it and so on, it is written clearly so the teacher actually knows what to do and what to achieve at the end of it. (Teacher E2)

Even though the tasks were designed with the syllabus in mind, there were some teachers who treated the tasks as enrichment tasks instead of non-routine activities for practice or problem solving.

> Advantages would be the girls are exposed to other kinds of activities... Activities are not exam based so it is something they can do for enrichment... The girls learn more about teamwork when they are doing this.

Some teachers believe that the tasks benefit some students more than the other in the class.

> It will benefit some of them, those who are really keen and those who are more motivated. (Teacher E1)

> I won’t say that it would benefit everybody... one thing for sure, it creates more enthusiasm… (Teacher E4)

Generally, the four teachers favoured the use of investigation tasks in the mathematics classrooms, albeit with some reservations.

2. Problems faced

All four teachers interviewed would like to use the tasks in the classroom but they seemed to have difficulties finding classroom time for the tasks. Apparently, they found it difficult to integrate the tasks into their scheme of work.

> I've always welcomed such (tasks)... but the only factor was the time factor. Sometimes what will be in my mind is that will I have time to carry it out. (Teacher E4)

> I find that these tasks need more time. Two periods is not enough to complete one task (Teacher E1)

> There is always the time constraint... (Teacher E2, E3)

Most investigations are open tasks and have different solutions. Consequently, it is time consuming to mark the investigations as pointed out by one of the teachers.

> You will take really a long time to mark cause if they are just drawing the lines everywhere (referring to a task that required students to locate all pairs of parallel lines and perpendicular lines formed with the tangram pieces), and there are so many lines also. So that will be very time consuming (to mark). (Teacher E3)

The other problem faced by the teachers in the task implementation is the poor language skills of some students in the class.

> some weaker students, they are unable to understand because it could be their command of English... (Teacher E2)

> some of them they can't read so they don't understand. (Teacher E4)

3. Teachers' Recommendations

According to the interviews, it was believed that to encourage the teachers to use more investigative tasks in class, more support from the Ministry would be needed. First, teachers must be given opportunities to be familiarized with both the tasks and the assessment rubrics which must not be too complex for the teachers to use. Supporting resource materials must be readily available as well because the thought of preparing the materials often discourages the teachers to use the tasks.

> We could have some little workshops where teachers get together to find out more about the activities first, try out on their own first, and figure out how they can actually improve on the lesson plans that were given. (Teacher E1)

> ... prepare the teachers (provide) introduction or preparatory kind of course... given them an idea. (Teacher E3)

> ... certain tasks requires manipulatives... it would be useful if they can have the actual manipulatives (instead of paper cuttings)... with that provided, would be very ideal. (Teacher E4)

In addition, the teachers expressed that they would prefer shorter investigation tasks with simple general rubric for assessment. The tasks may be made available in electronic form so that the teachers can modify them to suit their instructional objectives and the ability of their students. In addition, they would like the tasks to replace and not add on to some of the regular mathematics practise exercises.

One teacher pointed out that the effectiveness of the use of investigation tasks in the mathematics learning should not be tied to the performance in the school mathematics achievement tests. They are tangential advantages of using investigations in mathematics learning that are often not reflected in the regular achievement tests.

> To evaluate the effectiveness of the program based on students' result ultimately may not be fair to my personal opinion. (Teacher E4)

Student interviews
Two groups of three to four students were selected by each class teachers to participate in the interview. All students claimed that they like the tasks but they did have some difficulties with some of the task.

1. View of the investigation tasks

All the students pointed out that the investigation tasks given were more challenging and interesting than the tasks commonly found in the textbooks, activity books and worksheets. In addition, they had more fun carrying out the tasks, especially those involving games and items like chocolates. They believed that the tasks made them think harder. The tasks were also different because *"you have to write example, your feelings and your reflections."* One of the students pointed out that one had to go *"further than you are supposed to learn."*

However, students showed different views about the reflection component of the tasks. One of the students did not like it because "*I don't know how to explain our feelings."*

Others were aware of the benefits of writing the reflections. For example, one student said: "*We actually learn how to express what we want to say."* Another student indicated, "*The teacher can also find out what you are weak at."*

2. What they would like

According to the interviews, most students preferred simpler investigation tasks, which is quite understandable. They found some of the tasks too difficult and the task instructions too difficult to understand.

The interviews also revealed that all the students interviewed were aware of the importance of understanding the tasks and the role of teachers, that is, to explain and provide assistance when stuck.

In addition, all the students expressed that they would like to have more investigations, with the frequency suggested ranging from more than once a week to once in a month.

5.5 Conclusions and Recommendations

According to the information presented in this chapter, we can conclude the following for each of the research questions as mentioned at the beginning of this chapter.

(a) What is the impact of investigation/project work on the attitude of students towards mathematics?

While there are differences between groups (details are in Section 4), overall is seems reasonable to conclude that there was a general decline for all groups in the attitude to mathematics. The exception being the experimental low-ability group.

(b) What is the resulting attitude towards investigation/project work?

(i) *Students.* The majority of the students in each class seemed think that project work would be beneficial to be their mathematical development. Also, generally, the students like project work. Students like to do investigation tasks for various reasons, for example, the tasks are more challenging than the routine learning tasks, it is fun doing investigations However, this does not imply that there were no differences between the experimental classes. For example in class E3 only about half the students liked mathematics projects. Also, most students feel they need more guidance when going project work.

(ii) *Teachers*. The teachers were generally positive about the value of investigation/project work, but have some practical reservations. Some noted that all students who were keen and motivated would benefit from the use of investigation/project work.

(c) Does investigation/project work impact performance?

The results are mixed. On some items and in some classes the performance deteriorated from the pre-test to the post-test, while on other items and in other classes the performance improved. It also appears that students in School B had difficulty comprehending the question statements compared to the students in School A.

(d) What suggestions can be made for implementing investigation/project work in primary classrooms?

Teachers should capitalize on students' interest in investigation tasks and use the tasks to promote and assess learning.

Primary teachers are aware of the benefits of investigation tasks in mathematics learning. Hence, whether or not we should use investigation tasks as a mode to assess mathematics learning is not the issue here. The teachers would like to use the tasks in their classrooms but have some reservations. They need guidance and assistance from the relevant departments in the ministry. Generally, they would like to use investigation during instruction or as assessment of learning if the conditions for using the investigation are favourable.

Firstly, the tasks should "replace" and not "add on" to the already crowded primary mathematics curriculum. Investigations promote learning in depth. The current scope of the mathematics curriculum can be trimmed so that more curriculum time is available for students to investigate a topic in depth. Investigation involves higher order thinking and application of knowledge as well. Hence, it should be given during or soon after the instruction planned by the teacher for the same content. This is to ensure students have readily access to the necessary content of the investigation task. This would increase the effectiveness of investigations in concept / skill development and in the assessment of learning. Investigations are not meant to be used to engage students after the examination.

Secondly, relevant resources including manipulatives must be made accessible to teachers. Many teachers are discouraged to carry out investigations tasks in their classrooms because of the lack of manipulatives and the preparation of manipulatives is time consuming.

Lastly, most teachers do not know how to use the investigation tasks effectively. Professional development in the use of investigations in the mathematics classrooms must be provided for all teachers. Teachers need to be familiar with the tasks so that they can adapt the tasks for their own students. Besides the ability of their students, they also have to consider the gender of their students as there is the probability that the context of the investigation may favour the boys or the girls. In addition, they have to be trained in the design and use of assessment rubric to assess learning. Professional development in these areas can be provided by the National Institute of Education, the Curriculum Planning and Development Department and the Examination Board. Hence, to encourage teachers to use investigations in the mathematics classrooms, appropriate supports and incentives must be provided by both the Ministry of Education and the school administrators. The positive effects of investigations on mathematics learning as measured by the achievement tests are tangential. School administers must keep in mind that the effects on student learning cannot be readily reflected in the regular semestral examination.

Chapter 6 Results and Findings (IV): Communication Tasks (Primary)[7]

6.1 Introduction

As part of the Mathematics Assessment Project (MAP), this chapter focuses on the study of integrating communication tasks, more specifically journal writing, into primary mathematics learning is. As mentioned earlier, the MAP project focused on the integration of new assessment strategies into daily mathematics teaching and learning. Four new assessment strategies were researched in this project. They are communication-based assessment in mathematics learning, project-based assessment, self-assessment and performance-based assessment.

One of the main objectives in mathematics learning in Primary schools in Singapore is to develop students' ability to communicate their mathematical understanding. In fact, the Singapore mathematics syllabus for primary level emphasizes that students' ability 'to communicate mathematically...to be able to illustrate, to interpret, to explain and to discuss mathematical ideas and experiences in doing mathematics' (Ministry of Education [MOE], 2000, p. 17) is an important communication aspect in learning mathematics. In other words, part of mathematics learning is to develop students in using mathematical language to explain mathematical ideas and understanding precisely, concisely and logically. However, research has been done to show that in the primary mathematics classrooms, the textbook exercises and problems in the worksheets that students work on are usually structured in a way merely meant for drill and rote practice with not much sustained writing expected (Yeo & Zhu, 2005). Most of the time, worksheets are given to students for consolidation of topics learnt and students assimilate what teachers teach, in a seemingly procedural fashion. In addition, getting the correct answers seems to be the main indicator of students' understanding of mathematical sums and problems. Students should be given the opportunity to write substantially even in mathematics classrooms and it is believed that the use of alternative strategies such as journal writing could help in developing students' writing aspect.

In recent years, there has been an increasing interest in the use of journal writing as an alternative assessment tool in mathematics classrooms (Badger, 1992; Drake & Amspaugh, 1994; Pugalee, 2001). There are studies that attempt to classify and explain how students' writings in mathematics classes convey their understanding and learning of mathematics to teachers through the use of different writing activities (e.g. Pearce & Davison, 1988; Shield & Galbraith, 1998) and studies that assert the potential benefits of using writing to develop students' understanding of mathematics (e.g. Dougherty, 1996; Bagley & Gallenberger, 1992). Although the use of journal writing in mathematics learning is not prevalent in the context of our local schools, some small-scale structured research work in the mathematics classrooms such as Yazilah and Fan (2002) on journal writing with primary five students, Ng (2003) working with primary four to six and five students respectively in using journal writing as an alternative assessment, and Yeo's (2001) study in using journal writing in the junior college level seem to suggest the increasing recognition of developing students' mathematics learning through the use of writing.

With the recent rapid re-focus of how mathematics is taught and learnt in Singapore schools, there has been increased interest among Singapore teachers on how assessment could include alternative strategies to complement the traditional pen-and-paper assessment. This is to pave the way for a more holistic assessment of how mathematics is perceived and learnt at the primary school level. This is in line with Singapore's educational system to make necessary changes to overcome the inadequacies of the pen-and-paper assessments. This will definitely lend credence to the shift in values of education towards a more holistic and global approach.

This sub-study investigated how journal writing can be used to examine students' understanding of mathematics learning. The study focused on journal writing in mathematics as a means of communication in students' learning mainly through journal writing tasks. Journal writing tasks were incorporated into the Primary Three mathematics normal curriculum in two Singapore schools. The writing tasks specifically aimed to uncover how young students perceive the learning of mathematics.

It is hoped that this project can produce research-based evidence concerning the effectiveness of implementing journal writing in mathematics assessment in the Singapore educational setting.

6.2 Research Question & Conceptual Framework

6.2.1 Research questions

Similar to other components of the MAP project, the main research questions for this sub-study can be stated as follows:

1) What are the effects of using Journal writing strategies in the mathematics classrooms on students' achievement in both cognitive and affective domains?

2) How can journal writing strategies be effectively integrated in the teaching practice of mathematics teachers?

This study intended to investigate how journal writing can be used to examine Singapore Primary school students' perception and their ability to explain their understanding of mathematics learning.

A journal writing task is defined as a piece of work that a student writes through to explain and describe their knowledge, understanding and/or feelings, in learning of mathematics or about a particular mathematical concept or process, depending on the given writing prompts. It is hoped that journal writing will help bring out the communication aspect of the learning of mathematics. It is believed that through journal writing, students will be given the opportunity to express themselves and to connect cognitively and aesthetically with their teachers about mathematics. In addition, it is also believed that teachers, too, will be able to make better progress in understanding the young learners and hence enhance their own delivery of lessons.

The researchers believe that assessment, being an integral part of teaching and learning, can be incorporated into part of the daily work of students. Hence, in this study, the journal writing tasks are integrated into the normal tasks students attempt daily in the classroom. There is a need for teachers to understand their students' learning to be able to be more effective facilitators of the learning of mathematics in the young.

6.2.2 Conceptual framework

To understand deeper reasoning ability, teachers have to study the way students think. Sometimes, students think they have figured out an answer but their reasoning may not be quite right! The teacher's role then is to encourage their students to constantly examine their own thinking and then to help them find out and understand where their reasoning needs refining. Journal Writing in Mathematics learning helps both the teacher and the students themselves look at their learning. Integrating writing into the mathematics classroom can be easy for the teacher and beneficial for the students. Communicating about mathematics helps strengthen student learning, which can build deeper understanding. It provides students with an opportunity to organize their thoughts related to the mathematics topic, which helps clarify their thinking.

Student writing can also provide valuable insight for the teacher into their mastery of mathematics concepts. Teachers can use writing assignments as either an informal or formal assessment tool. Writing often reveals gaps in learning and misconceptions, which can help inform the teacher's instructional planning and intervention strategy.

Communications in Mathematics learning is one of the objectives of mathematics learning in Singapore. Undoubtedly, pen-and-paper test has been the main assessment mode in the Singapore School context. However, the traditional pen-and-paper tests do not sufficiently bring out the mathematics communication skills in students. Journal writing complements the inadequacies of the pen-and-paper assessment.

Participating students were taken through the 3 basic types of journal prompts, classified in this sub-study.

1) *The content prompt*, which focuses on what students know of a particular concept. An example of a content prompt is when students are asked specifically to describe what they think "4 x 6" means to them.

2) *The process prompt*, which helps teachers understand the process of students' thinking through a concept. An example of a process prompt given is when students are asked how they will measure and cut a length of ribbon to a specific length using 2 other pieces of strings of different lengths as the measurement tools.

3) *The affective prompt*, which helps teachers gain a better overview of how students feel at certain points in the learning of mathematics or of a particular mathematical event. An example of an affective prompt given to the students asks students how they feel about how they have done in a recent mathematics pen-and-paper examination.

Because of the nature of the study where teachers were encouraged to integrate journal writing into their everyday tasks, the journal writing prompts were designed to vary according to what the teachers would want for their own class of students.

6.3 Methods

6.3.1 Participants

As discussed earlier, eight Secondary and eight Primary schools were involved in this large-scale project on integrating new assessment strategies into daily mathematics classroom teaching and learning. In particular, two primary schools participated in this research on the use of journal writing in mathematics learning in the primary school. One is a high-performing school and the other, a non-high performing school. In order to have more representative samples from each school, two classes of Primary Three students were selected such that one had high-ability (HA) and the other low-ability (LA) students. A total of four classes or 158 students and four teachers were involved in this experimental study. The school that was identified as a high performing school ranked as one of the top nine best performing schools based on the Primary School Leaving Examination Results in year 2003. Comparison classes were matched to the experimental classes one-on-one as far as possible.

The participating teachers were trained teachers. In the first year of the study, the teachers in the high performing school experimental classes had an average of 15 years' experience. Those in the non-high performing school experimental classes had an average of 5.5 years of teaching experience.

Because of some unforeseen reasons and the schools' plan, all 4 first-year participating teachers were unable to continue with the project into the second year. Four new teachers replaced the previous participating teachers to continue with the project in the second year of MAP. These participating teachers in the 2nd year of study were also trained teachers, averaging 12 years' experience in the high-performing school and 3.5 years in the non-high performing school. The comparison classes were selected based on their compatibility in terms of academic performance. Teachers in the high performing school comparison classes for the 1st year averaged 8 years of teaching experience and for the 2nd year average 27.5 years. Teachers in the non-high performing school comparison classes for the 1st year averaged 5.5 years of teaching experience while those for the 2nd year also averaged 5.5 years.

A profile of these teachers in terms of their professional experiences is summarized in Table 6.1. Due to practical reasons, it was impossible for the research team to select comparison classes strictly equivalent to experimental classes for both teachers and students.

Table 6.1. Average no. of years of teaching experiences of teachers in experimental and comparison classes

	High-performing School		Non-high-performing School	
	1st year (2004)	2nd year (2005)	1st year (2004)	2nd year (2005)
Experimental classes	15	12	5.5	3.5
Comparison classes	8	27.5	5.5	5.5

More specific information about the participating students and teachers' particulars are presented below:

Experimental classes in the high-performing school
At the end of 2004 (Y2004), seven students in the HA class left the class and the participating teacher was not able to carry on with the project for another year. Six new students and a new teacher joined the HA class in Y2005.

Four students left the LA class in the year ending Y2004 and five new students joined the LA class in Y2005. A new teacher replaced the previous participating teacher who was unable to continue with the study.

Table 6.2 presents the detailed information about the participating classes (here including comparison classes) in the high-performing school, while Table 6.3 shows the information about the participating teachers (here including teachers of comparison classes) in that school.

Table 6.2. Participating classes in the high-performing school

2004 (1st year)		2005 (2nd year)	
Type of class	No. of students	Type of class	No. of students
HA (High Ability)	42 – experimental	HA (High Ability)	41 – experimental
	40 – comparison		41 – comparison
LA (Low Ability)	36 – experimental	LA (Low Ability)	37 – experimental
	36 – comparison		37 – comparison

Table 6.3. Participating teachers in the high-performing school

Experimental Group (1st year) 2004			Experimental Group (2nd year) 2005		
Type of class	Teacher	Years of teaching	Type of class	Teacher	Years of teaching
HA (High Ability)	HH1	9	HA (High Ability)	HH2	20
LA (Low Ability)	HL1	21	LA (Low Ability)	HL2	4.5
Comparison Group (1st year) 2004			**Comparison Group (2nd year) 2005**		
Type of class	Teacher	Years of teaching	Type of class	Teacher	Years of teaching
HA (High Ability)	HH3	6.5	HA (High Ability)	HH4	14
LA (Low Ability)	HL3	10	LA (Low Ability)	HL4	41

Experimental classes in the non-high-performing school
In the experimental HA class, all students moved up en masse to a new class in Y2005. However, a new teacher took over as the previous one has left the school.

At the end of Y2004, two students left the LA class and two new inclusions were made in Y2005. A new teacher replaced the previous teacher who was unable to continue with the project.

Table 6.4 presents the information about the participating classes (including comparison classes) in the non-high-performing school, while Table 6.5 details the information about the participating teachers (here including teachers of comparison classes) in the school.

Table 6.4. Participating classes in the non-high-performing school

2004 (1st year)		2005 (2nd year)	
Type of class	No. of students	Type of class	No. of students
HA (High Ability)	40 – experimental	HA (High Ability)	40 – experimental
	40 – comparison		40 – comparison
LA (Low Ability)	40– experimental	LA (Low Ability)	40 – experimental
	40 – comparison		40 – comparison

Table 6.5. Participating teachers in the non-high-performing school

Experimental Group (1st year) 2004			Experimental Group (2nd year) 2005		
Type of class	Teacher	Years of teaching	Type of class	Teacher	Years of teaching
HA (High Ability)	NHH1	5.5	HA (High Ability)	NHH2	1.5
LA (Low Ability)	NHL1	5.5	LA (Low Ability)	NHL2	5.5
Comparison Group (1st year) 2004			**Comparison Group (2nd year) 2005**		
Type of class	Teacher	Years of teaching	Type of class	Teacher	Years of teaching
HA (High Ability)	NHH3	5.5	HA (High Ability)	NHH4	4.5
LA (Low Ability)	NHL3	5.5	LA (Low Ability)	NHL4	6.5

6.3.2 Instruments

Similar to other components of the MAP projects, a number of research instruments were designed and used to collect the data for this sub-project, which are described below.

Questionnaires
A pre-survey questionnaire was designed for the researchers to understand the students' attitudes and dispositions towards mathematics and the learning of mathematics. The survey was structured into Parts A and B.

Part A, consisting of 14 items, was administered to all sub-groups of the MAP. The 14 items were designed as follows:

5 items pertaining to general views on the learning of mathematics,
Qn 1: I enjoy doing mathematics.
Qn 6: Mathematics is useful to me.
Qn 7: Mathematics is hard for me.
Qn 12: I dislike mathematics.
Qn 14: I like spending time on studying mathematics

3 items on anxiety issues,
Qn 2: I am afraid of doing mathematics
Qn 9: I don't feel good toward mathematics
Qn 10: It makes me nervous to do mathematics

4 items on perception of own performance in mathematics
Qn 3: I am sure I can do mathematics well.
Qn 4: I can get good grades in mathematics.
Qn 11: I am not good at mathematics
Qn 13: I like to do difficult mathematical questions.

2 items on their beliefs in the usefulness of mathematics
Qn 5: I think mathematics is useful to me.
Qn 8: Knowing mathematics will help me get a good job next time.

Part B consisted of 9 questions asking the frequency different mathematical tasks were given by the teacher as well as types of questions with more than one answer and those that were strictly of no consequence in daily life application. Of the 9 questions which were given to all participants for the large-scale MAP, 2 pertained to this particular study of journal writing in mathematics.

- Qn 15C: my math teacher had asked me to explain mathematics ideas in writing.
- Qn 15D: my math teacher had asked me to write down the reasons for my math answers.

After eighteen months into the research project, *a* post-survey was conducted. For the experimental classes, Part A of the pre-survey questionnaire was replicated in Part A of the post-survey. But for Part B, there was a slight change in the questionnaire. 15 questions pertaining only to journal writing were included to uncover insights into the participants' thinking. As the comparison class students were not exposed to the use of journal writing in their learning of mathematics, Parts A and B of the pre-survey items were given to them in the post-survey questionnaire.

Part B for experimental classes includes the following items:

- Qn 15: I like to write mathematics journals.
- Qn 16: Writing mathematics journals is easy to me.
- Qn 17: Writing mathematics journals helps me to learn mathematics.
- Qn 18: I am not afraid of writing mathematics journals.
- Qn 19: Writing mathematics journals helps me to be more aware off my understanding of mathematics.
- Qn 20: Writing mathematics journals is an important skill in mathematics learning.
- Qn 21: I am able to express about my feeling toward mathematics through writing mathematics journals.
- Qn 22: I am able to tell others about my understanding of mathematics through writing mathematics journals.
- Qn 23: I don't know how to get started when I am writing mathematics journals.
- Qn 24: Writing mathematics journals makes me think broader and deeper about mathematics.
- Qn 25: When I am writing mathematics journals, I know what I am expected to write.
- Qn 26: I can write my mathematics journals well.
- Qn 27: I would like to have more mathematics journal writing for my mathematics lessons.
- Qn 28: Writing mathematics journals makes me learn mathematics better.
- Qn 29: Writing mathematics journals is a waste of time.

New strategy tests
The pre-test was designed taking into consideration that multiplication is the major emphasis in the Primary Three syllabus.

In fact, multiplication and division form the main topics of mathematics learning at Primary Three level. A large portion of curriculum time was apportioned to the learning of Multiplication and Division at this level. Moreover, multiplication and division are important concepts in mathematics with very broad application purposes in our daily lives. A lot of interest has been generated concerning the interpretation of a multiplication sentence, i.e. three numbers with the symbol "X" in their relationships.

Below is the journal prompt given to 158 participants from the two participating schools.

PRE-TEST *You have learnt about multiplication. A new pupil is going to join you and your teacher wants you to explain to your new friend what this means:* *6 x 4 = 24*

You can use pictures to explain too. Write down how you are going to go about it. Tell as much as you can about 6 x 4 = 24.
Remember you ARE the teacher now.

The post-test is parallel to the pre-test. Teachers were consulted and they felt that the results would help them understand if the students were able to communicate better after the exposure given in the form of intervention tasks.

POST-TEST

You have learnt quite a lot about multiplication. A new pupil is going to join you and he has not learnt multiplication yet. Your teacher wants you to explain to this new friend what this means:

7 x 5 = 35

You can use pictures to explain too, if you think it will help your friend.
Write down how you are going to go about it. Write as much as you can about 7 x 5 = 35.
Remember **you ARE** the teacher now.

Intervention tasks
The researcher started with the intervention tasks. They were mainly journal prompts in print. Discussions were carried out between the researcher and the teacher participants. Participating teachers were given a short workshop by the central team to familiarize with and had a common understanding of what journal writing in mathematics is about. Subsequently, teacher participants volunteered to craft their own journal prompts. In total, more than 50 intervention tasks had been carried out to give students the platform and avenue for expressing both their cognitive learning of mathematics as well as their disposition towards mathematics learning.

Regular meetings about once a month were held between the researcher and the teacher participants besides the observations. During the meetings, researcher and teacher participants discussed issues pertaining to the tasks. However, there were time constraints to effect a better communication between the parties involved.

Tables 6.6 and 6.7 give a list of the interventions carried out, complete with the date and topics in the high-performing and non-high-performing schools respectively.

Table 6.6. A list of intervention tasks implemented in the high-performing school

Task no.	Type of class	Date	Topic
1	HA	25 Mar 04	Long Multiplication
2	HA	15 Apr 04	Long Division 1
3	HA	22 Apr 04	Long Division 2
4	HA	29 Apr 04	Tables and Columns heuristic
5	HA	20 May 04	Reflecting SA1 results
6	HA	8 Jul 04	Patterns (MI)
7	HA	29 Jul 04	Money: The price is right
8	HA	5 Aug 04	Length
9	HA	20 Jan 05	Maths as a colour
10	HA	11 Feb 05	Estimation: why learning it is so important
11	HA	16 Feb 05	Estimation: How you can use it

12	HA	24 Mar 05	Geometry: writing a poem with properties
13	HA	3 Mar 05	Geometry: how to use the protractor
14	HA	7 Apr 05	Fractions: Affective
15	HA	21 Apr 05	Multiplication tables: Look for patterns
16	LA	25 Mar 04	Multiplication and Division
17	LA	1 Apr 04	Division 2
18	LA	29 Apr 04	Tables and Columns heuristic
19	LA	19 Jul 04	Money: the price is right
20	LA	22 Jul 04	Length
21	LA	12 Aug 04	Length
22	LA	16 Feb 05	Maths as a colour
23	LA	1 Mar 05	Angles
24	LA	24 Mar 05	Magic Triangle
25	LA	15 Apr 05	Fractions
26	LA	21 Apr 05	Folding fractions

Table 6.7. A list of intervention tasks implemented in the non-high-performing school

Task No.	Type of Class	Date	Topic
1	HA	11 Mar 04	Multiplication : Pre-test
2	HA	2 Apr 04	Division (50 divided by 5)
3	HA	13 Apr 04	Division
4	HA	30 Apr 04	Long Division
5	HA	1 Jul 04	Mental Calculations
6	HA	13 Jul 04	Length with score sheet / 1 questionnaire
7	HA	19 Aug	Graph
8	HA	3 Sep	Fractions
9	HA	27 Sep 04	Geometry
10	HA	20 Jan 05	Whole numbers
11	HA	4 Feb	Affective : I want to be better at Maths so that …
12	HA	3 Mar 05	Multiplication by 2d
13	LA	11 Mar 04	Multiplication : Pre-test
14	LA	2 Apr 04	Division (50 divided by 5)
15	LA	13 Apr 04	Division
16	LA	30 Apr 04	Long Division
17	LA	1 Jul 04	Mental Calculations
18	LA	13 Jul 04	Length with score sheet / 1 questionnaire
19	LA	20 Aug	Graph
20	LA	2 Sep	Fractions
21	LA	27 Sep 04	Geometry
22	LA	20 Jan 05	Addition of whole numbers
23	LA	27 Jan 05	Nos to 100 000
24	LA	3 Feb 05	Affective: I want to be better at Maths so that …
25	LA	3 Mar 05	Explain 3d x 2d sum

Classroom observation (including video taping)
The researcher conducted classroom observations of the teachers conducting intervention measures. It was hoped that the teachers could be captured actively conducting the intervention measures. This included video-record some lessons with the permission of the participating teachers.

During the observation, the researcher did not interfere at all in the teaching and learning of mathematics in the classrooms. However, to video-record and observe at the same time proved to be a big challenge for the researcher. Students were initially curious about the presence and actions of the researcher but were quick to adapt to the subsequent visits by the researcher. A total of 10 and 9 classroom observations / video recordings for the high performing and non-high performing schools respectively were carried out.

Interviews
Towards the completion of the intervention of the project, interviews were conducted. Student and teacher-participant interviews were conducted for only the experimental classes. A total of 24 students from both the high-performing and the non-high-performing schools in the experimental group were interviewed. In each of the participating schools, 12 students were selected by the teacher participants for the interview by the researcher. Details are shown in Table 6.8 below.

Table 6.8. Teachers and students interviewed

Schools	Experimental HA class	Experimental LA class
High-performing	6 students and 1 current teacher	6 students and 1 current teacher
Non-high-performing	6 students and 1 current teacher	6 students and 1 current teacher

The central research team crafted the interview questions. The interview questions aimed to uncover how participants felt towards this new assessment on journal writing in mathematics. They were designed to obtain feedback from the participating teachers and students on what had been carried out as intervention measures and on views of the participants on the usefulness of the new assessment strategy. The questions also included views from participants on the feasibility and concerns in implementing the new assessment strategy and the preferred frequency should this mode of assessment be implemented nationwide.

Student interviewees were interviewed in groups of 3 at any one time. Participating teachers in the experimental groups were interviewed one at a time. As mentioned, the teacher interviewees were the “current” new teachers who took over the experimental classes in the beginning of 2005. Teacher-participants who had been with the project in the year 2004 were not available for the interviews. The teacher interviewees, hence, were with the project for about six months at the time of the interviews. All the interviews were recorded and then transcribed.

6.3.3 Procedures and data collection

Some relevant information about the procedure and data collection has been mentioned earlier, and hence will not be repeated in this section.

The pre- and post-tests were designed by the researcher who was in charge of this sub-project, with the consultation with some other team members, especially the principle investigators of the project. They were administered to 4 experimental as well as 4 comparison classes. In total, 8 classes of about 40 students each took part in the tests. They included both experimental and comparison class participants. The participants in this study were all 9 year-olds in Primary Three in the first year and a similar number who continued as Primary Four students in the second year.

There was a slight change in the candidates for the post-tests. The participating students were slightly older, having gone to a level higher in Primary Four. The change in numbers was negligible. The post-test was designed in consultation with participating teachers. The tests were administered by the respective participating teachers on the same day although the timing may be slightly different. Students perform the tests under examination conditions. However, a little flexibility was allowed to the duration of time spent on the task.

The pre-test was carried out in March 2004. The post test was administered in May 2005. These tests were graded by the researcher using the rubrics designed for the test items. Most intervention tasks were prepared by the participating teachers citing it to be more effective as they can be done closely with the Schemes of work of the school departments and the syllabus. Teachers also preferred to do the marking themselves. Classroom observations were decided on between the researcher and the teacher-participants. These observations were conducted with the consent of the teacher participants. The researcher collected all data from participating teachers. Intervention tasks were kept by the researcher. Results from surveys and tests were collected by the researcher who then submitted them to the central team for analysis. Transcripts to interviews were outsourced.

6.3.4 Limitations of the study

After the schools were selected, the school leaders volunteered their schools to be part of the study. Teachers teaching the targeted classes were naturally expected to take on the challenge. Teachers were in a sense not really given a choice. In other words, they to a large degree was requested to participate in the research study. It took some convincing for the researchers to enthuse them into taking on the task seriously and become active participating teachers. Fortunately, most were cooperative and tried to do their best after being convinced of the benefits of the study.

In the second year of the study, all participating teachers did not follow up their classes. This was something beyond the control of the researchers. The school had its own needs and staff deployment was done by the school leaders. The study carried on with the new group of participating teachers. Again, these teachers were essentially not given a choice. They took on the roles as teachers of the participating classes. As these teachers were new to the class, their personalities and attitudes did make a difference in the behaviour and performance of the students in the study.

The choice of teachers, some positive, some proactive, some indifferent and apprehensive to some degree did affect the research findings and results. The short duration of time in which the second year teachers had in acclimatising themselves to the research study also affected the rapport with the students in question.

Time was also a crucial issue. Having to juggle formal work arrangements outside of the research study with observations, talk-times with teachers, data collection, report writing and task formation is a real challenge. However, members of the central MAP team had been very forthcoming in extending their help when requested.

It is good that the post-test was a parallel one. However, because of the students' slight maturity after a year, they seemed to lose interest in what they called a “familiar” or “boring” task.

The observations cum video-taping sessions were a real challenge to the researcher. Trying to observe closely the nuances taking place in the classroom was rather compromised in the earnest hope of capturing the intervention measures on video. Furthermore, the insufficient number of video cameras added to the inconvenience of having to loan from the participating schools. The schools also had limited numbers. The situation could be quite bad when the researcher had to grapple with operating the schools' video cameras without necessary help. Although eventually the researcher was able to make out how to operate them, some good moments in the classroom were lost.

6.4 Results

6.4.1 Findings about students' attitudes

We first report the results from the high-performing school.

Attitudes towards mathematics and the learning of mathematics
The data from questionnaire surveys revealed that students in the experimental/high ability class provided more negative responses in the post- than pre-survey on all the items and on all but 1 item (Item 8) they provided significantly more negative responses in the post-survey; In contrast, students in comparison/high class provided more positive responses in the post- than pre-survey on all but 7 items (Item 3, Item 4, Item 7, Item 9, Item 11, Item 13, Item 14) but no significant difference was detected. The results also showed students in comparison/high class were more negative in the perception of their own performance in mathematics.

The result seems not unexpected to us. Students of this experimental/high ability class were extremely active and individualistic in delivering their tasks. Classroom management is a challenge. The 1st year participating teacher had to spend a lot of time before and during tasks to keep the students focused and engaged. The 2nd year participating teacher was more didactic in her approach although she did have more creative ideas in terms of journal tasks. Her diligent administration of tasks may not have appealed to this very active class who were constantly challenging her to the tasks given. We believe that other factors including particularly teachers' teaching approaches played a more important role in this aspect. Moreover, the result suggests that using journal writing alone is not enough for resolving the issue about students' attitudes towards mathematics and the learning of mathematics

Similarly, students in experimental/low ability class provided more negative responses in the post- than pre-survey on all the items and on 1 item (Item 13) they provided significantly more negative responses in the post-survey; In contrast, students in comparison/low class provided more positive responses in the post- than pre-survey on all but 6 items (Item 7, Item 9, Item 10, Item 11, Item 12, Item 13) but no significant difference was detected;

The 1st year participating teacher was not able to continue being with the target group of students into the second year. A different teacher took over the research tasks. As this 2nd year teacher was experiencing a problematic pregnancy, she registered a high absentee rate. Relief teachers taking over the lessons were not able to carry out the tasks. The rapport with her students was also adversely affected by her periods of absence. The students seemed to lack direction and drive in their work as noticed during the classroom observations made by the researcher.

In the interview with the teacher, she indicated she accepted the project as she was given no option. So she was resigned to it and tried to keep an open mind about the MAP although she felt strongly that the journal tasks *"should not replace the "regular mathematics exercises" although the regular mathematics exercises "are monotonous... but we still need to go through.."* The teacher also mentioned that she thought only *"20% of her students are actually interested in Journal writing."* Below is from interview.

> Researcher: To someone new to journal writing in Maths, what kind of advice will you give to him/her?
>
> Teacher: I think he should go with an open mind because when I first started I was a bit apprehensive... since I have to do this, might as well give a shot.

> Researcher: ... do you think it is appropriate to replace some regular math exercises with this journal writing? How feasible do you think it is?
>
> Teacher: I don't think there should be any replacement for any of the others. I feel that it is just an add-on. The others although are monotonous, but it's still I think what we need the children to go through. For journal writing, it is just another outlet for expression.

From the interviews, students mentioned that although they saw the benefits of journal writing, they found that they "*have to explain why, how to do this"* which workbook exercises did not include.

> S: It's like you have to write for Maths journal and then because it's Maths journal, you have to explain why, how to do this...

In the interview conducted, students seemed to find "it very hard, usually we don't like to explain things like that". They also found that sometimes the journal tasks were very boring, specially citing the post-test, "But sometimes it quite boring because we keep giving the same sums like multiplication." The teacher felt that she was not experienced enough to craft good journal prompts, resulting in the students finding the intervention tasks "repetitive" and "boring".

In the pre-survey, students in experimental/high ability class provided more positive responses than those in comparison/high ability class on all the items and on 4 items (Item 1, Item 2, Item 7, Item 11) they provided significantly more positive responses; In contrast, in the post-survey, students in experimental/high ability class provided more negative responses than those in comparison/high ability class on all the items and on 5 items (Item 1, Item 3, Item 5, Item 9, Item 10) they provided significantly more negative responses;

There was quite a marked difference in teacher presentation of lessons. The 1st year teacher in the experimental/high ability class was more tolerant of mistakes and try-outs. The 2nd year teacher, being a very experienced teacher, has definitely higher expectations of the students. The students were not used to her didactic approach and as a result are not co-operative in their learning. This could have impacted their self-esteem, resulting in the way they view mathematics as most primary school students would associate their liking of the subject with the teacher. Could the students have found it repetitive and boring to do similar tasks throughout the research period contributed to the more significantly negative responses in the post-tests? If so, what is the reasonable frequency? This is something remaining to be further studied.

Similarly, in the pre-survey, students in experimental/low ability class provided more positive responses than those in comparison/low ability class on all but 4 items (Item 4, 8, 11, 12) but no significant difference was detected. In contrast, in the post-survey students in experimental/low class provided more negative responses than those in comparison/low class on all the items but no significant difference was detected. Again, we think that other factors including particularly teachers' teaching approaches might have played a more important role in this aspect. The function of journal writing cannot be overestimated in promoting students' attitudes towards mathematics and the learning of mathematics.

Attitudes towards communication tasks
As said earlier, the results are from Part B in the post-survey, containing 15 items, conducted in the experimental classes. The data showed that students provided overall positive responses to the majority of the items; in particular, it was the case for all but 1

item (Item 27) for experimental/high ability class and all but 3 items (Item 15, Item 23, Item 27) for experimental/low ability class.

Interviews with the experimental/high ability class showed that all the student interviewees agreed that journal writing made them think and helped them in learning mathematics better. In addition, all of them wanted to continue journal writing although they said that some of their classmates thought otherwise (Refer to transcripts below. T: researcher; S: students).

> (Transcripts for experimental/high ability class)
> T: Would you give support for journal writing to continue in your class?
> S1/S2/S3: Ya/Yes/Yes.
> T: Ok. What about your classmates? Do you think your classmates would welcome the idea of having journal writing in class?
> S2: I don't think so.
> T: Why not?
> S2: They think it's very, a waste of time where you have to think.
> S1: Yes. They might not want the journal writing because they think it's a waste of time like that.
> S3: I think some of them want coz' some of them appreciate it. They think it's fun.
> T: If you were given a choice, would you like journal writing to be continued in class?
> S4: Mm...Yeah but I would like to have journal writing once a month or once a fortnight. Because I think if we have too many journal writings, we will somehow mix up with our school work or tuition work and we will have too much work to do.
> T: Do you think you want journal writing to be part of class work?
> S5: Yes …because if we don't do it, the teacher may not think that we understand or we copy from our friends. So if we do our journal individually, then the teacher knows we understand. If we don't, she can explain to us again.
> S6: Yes, because journal writing is easier than school work.
>
> (Transcripts for experimental/low ability class)
> T: Ok thank you. Now, do you support more journal writing in class?
> S7: Yes I do, because I find it's very interesting and it's actually not a waste of time like the others.
> T: So you say that you will support having more journal writings in class.
> S7: Yes I will, because I found them very interesting and not a waste of time like other things.
> T: What do you mean by other things are a waste of time?
> S7: I mean other children in class. They think it's a waste of time because they rather spend their time doing something else.
> T: Ok. What about you? Would you support the use of journal writing?
> S8: Ya because it's like more challenging and you will know more knowledge and you will learn more methods.
> T: Ok thank you. What about you? A waste of time? It's ok to speak your mind.
> S9: Ya sometimes it's interesting. But sometimes it's quite boring because we keep giving the same sums like multiplication.
> T: Ok. Now, what do you think about your classmates? Do you think your classmates would want to have more journal writing?
> S7: No, I don't think so. That's because they seem very bored and are only interested when they can do group work. So they will wreck their brains to get out of one problem till they get the answer. And feel satisfied just lean back and relax.
> T: Ok. What led you to say this? You must have heard them talking about it.
> S9: Ya, they never like Maths journals. They always like very fast do and then they can go and do their own things.
> T: Ok. What about you?

S8: They always like finish very fast and then they'll start playing and then if they don't play, they'll start reading comic book and then the teacher will confiscate and then when they get back their papers, they'll get wrong.
T: Ok. Now, I'm going to ask your own opinion. Would you like journal writing in a Mathematics class?
S10: Yes.
T: Why do you find it so?
S10: Because I find it a little interesting.
T: What about you?
S11: Yeah, I want to continue.
T: Mm, why would you want journal writing to continue?
S11: Because sometimes it's interesting to get to do it in pairs. Because someone in the group wants to do it and then you can ask someone and then you can relax while we do it.
T: Mm, ok. What about you?
S12: Yeah.
T: You think you like journal writing?
S12: Yeah, we can also learn more things.
T: Mm. Ok, now let's say the teacher tells you we're going to have journal writing in the class. Now the teacher wants your opinion to decide. How often do you want journal writing to be?
S10: Like er 5 times a month.
T: 5 times a month. What about you?
S11: Maybe 2 times a week.
T: Ok, how about you? How often do you want it?
S12: Once a week.
T: Why once a week?
S12: Because if you do it many times, it will seem like very stressed.

Students in experimental/high ability class provided more negative responses than those in experimental/low ability class on all but 7 items (Item 15, Item 16, Item 18, Item 21, Item 22, Item 23, Item 26) but no significant difference was detected.

(Transcripts for experimental/high ability class)
T: Ok. Now, what are the benefits? Do you find that journal writing actually helps you in learning Mathematics better?
S1: No.
S2: A little. It teaches you to explain about more and if anybody needs help you can explain more to them.
T: So Shaun, you think it does help?
S2: Ya, a little.
T: A little. Ok. What about you, what do you think?
S1: Ok.
T: What about you?
S3: Help us to learn how to teach. So if you grow up you become a teacher then you know how.
T: Oh, so you want to be a teacher?
S3: No.
T: Ok. Now, do you think this journal writing helps you in your Maths?
W: Yes.
T: Do you always think so?
S4: Yes, I think so.
T: How come you think so? What makes you think so?
S4: It helps us to em… regain what our teachers have taught us.
T: M hm. What about you?

S5: Because it helps us to understand more about the fractions and all the other subjects and other things.
T: Thank you. What about you?
S6: I just feel like it helps us but then…
T: M hm.
S6: Em, these things… ah I forgot what I want to say.
T: It's ok (said laughingly). Now what are the things you actually learn from doing journal writing? Did you learn anything new with journal writing as compared to other daily work? Anybody's ready to talk?
S4: It helps us to express our feelings and to learn deeper in maths.
T: Ok, any other comments?
S5: Mm. Not really. It doesn't really ask because it just helps us understand more.
T: Ok, good. Now, do you think journal writing has something negative you don't like?
S4/S5/S6: No.

Now let us turn to the non-high-performing school about the results from the surveys.

Attitudes towards mathematics and learning of mathematics:
Students in experimental/high ability class provided more negative responses in the post- than pre-survey on all but 2 items (Item 5, Item 13) and on 2 items (Item 12 Item 14) they provided significantly more negative responses in the post-survey; Similarly, students in comparison/high class provided more negative responses in the post- than pre-survey on all the items and on 2 items (Item 9, Item 10) they provided significantly more negative responses in the post-survey.

Students in experimental/low ability class provided more negative responses in the post- than pre-survey on all the items and on all but 2 items (Item 9, Item 10) they provided significantly more negative responses in the post-survey; In contrast, students in comparison/low class provided more positive responses in the post- than pre-survey on all but 1 item (Item 10) and on 1 item (Item 12) they provided significantly more positive responses in the post-survey;

In both the pre- and post-survey, students in experimental/high class provided more negative responses than those in comparison/high class; in particular, it was the case for all but 1 item (Item 10) in the pre-survey and all but 2 items (Item 1, Item 12) in the post-survey; moreover, while students in experimental/high class provided significantly more negative responses than those in comparison/high class on 5 items (Item 4, Item 5, Item 6, Item 8, Item 13) in the pre-survey and no significant difference was detected in the post-survey;

In the pre-survey, students in experimental/low ability class provided more positive responses than those in comparison/low class on all but 1 item (Item 11) but no significant difference was detected; In contrast, in the post-survey, students in experimental/low class provided more negative responses than those in comparison/low class on all but 3 items (Item 2, Item 3, Item 10) and on 1 item (Item 14) they provided significantly more negative responses. It seems to us that the interpretations offered earlier can be similarly used to explain these results here.

Attitudes towards communication tasks
Again, the results were from Part B in the post-survey, containing 15 items, conducted in the experimental classes.

The results showed that students provided overall positive responses to the majority of the items in this part; in particular, it was the case for all the items for experimental/high class and all but 2 items (Item 16, Item 26) for experimental/low class.

Students in experimental/high ability class provided more positive responses than those in experimental/low class on all the items and on 7 items (Item 15, Item 16, Item 18, Item 21, Item 22, Item 25, Item 16) they provided significantly more positive responses. Overall, the results suggest that students can accept journal writing tasks.

6.4.2 Findings from the intervention

As mentioned earlier, the participating teachers integrated the journal writing tasks into the normal daily teaching curriculum and the journal writing tasks were completed during curriculum time. In fact, the journal writing tasks were infused into students' daily learning as part of their everyday tasks they would have attempted. On completion, students' journal scripts were collected and marked by participating teachers.

An analysis of students' writings reveals that the use of appropriate journal writing prompts gave students the opportunity to explain their understanding and perception of mathematics learnt. The results from students' journal writing helped teachers to understand their students' thinking and take the necessary measures to clarify the mathematical concepts involved in the topics taught. Moreover, students' writings served as a source of reflection for teachers to re-think and improve their instructional strategies.

The results also suggested that teachers can gather a lot of information about students' thinking in their writings and journal writing can be an effective and alternative way to increase teachers' understanding of their students' learning in mathematics. Given the appropriate writing prompts, journal writing can be a source for students to reveal to teachers their understanding of mathematics concepts and learning. Nevertheless, the students' journal writing entries also showed that students need more exposure and opportunity to explain mathematics explicitly and substantially in writing.

Overall, the researchers believe that through writing, students will learn to develop and foster a higher ability to communicate their mathematical ideas and understanding.

6.4.3 Results from pre- and post-tests

A comparison of the students' performance in the post-test shows an improvement in the students' ability to express their mathematical understanding. However, for the high-performing school HA class, the students appeared to lack the enthusiasm and stamina to want to explain even better.

In the high-performing school experimental /HA class, only 43.9% gave explanations that included more than one way of presenting as compared with 45% in the pre-test. The experimental/LA class improved in the same category from 19.4% to 23.5%. The comparison/HA class improved from 2.6% to 17.9% and the comparison/LA class also improved from 17.1% to 27%.

Table 6.9 and Table 6.10 presents more detailed information on the students' performance in the pre- and post- tests on new assessment strategy tasks in both the experimental and comparison classes in the high-performing school and non-high-performing school, respectively.

Table 6.9. Students' performance in pre- and post-tests on new assessment strategy tasks (High-performing School)

Pre-test	*Experimental Class*				*Comparison Class*			
	3 Boldness (HA) 42 (2 absent)		3 Initiative (LA) 36 (all in)		3 Attentiveness (HA) 40 (1 absent)		3 Joyfulness (LA) 36 (1 absent)	
	No.	%	No	%	No.	%	No.	%
6 groups of 4	20	50	22	61.1	26	66.6	24	68.6
4 groups of 6	2	5	6	16.7	9	23.1	2	5.7
Both 6 groups of 4 and 4 groups of 6	18	45	7	19.4	1	2.6	6	17.1
Incoherent (others)	0	0	1	2.8	3	7.7	3	8.6
Total	40	100	36	100	39	100	35	100
Post-test	*Experimental Class*				*Comparison Class*			
	4 Boldness (HA) 41		4 Initiative (LA) 37 (3 absent)		4 Attentiveness (HA) 41 (2 absent)		4 Joyfulness (LA) 41 (4 absent)	
	No.	%	No	%	No.	%	No.	%
7 groups of 5	7	17.1	10	29.4	25	64.2	23	62.2
5 groups of 7	15	36.6	9	26.5	7	17.9	4	10.8
Both 7 groups of 5 and 5 groups of 7	18	43.9	8	23.5	7	17.9	10	27.0
Incoherent (others)	1	2.4	6	17.6	0	0	0	0
Total	41	100	34	100	39	100	37	100

In the *non-high-performing school experimental/HA class,* 15% give explanations that included more than one way of presenting in the pre-test. Comparatively, in the post-test, 53.8% of this group of students was able to do so. The *experimental/LA class* improved in the same category from 30.8% to 47.0%. The *comparison/HA class* improved from 38.5% to 62.2% and the *comparison/LA class* also improved from 19.4% to 63.2%.

The students in the non-high-performing school appeared to write and explain more. Enthusiasm in the post-test was apparent.

Table 6.10. Students' performance in pre- and post-tests on new assessment strategy tasks (Non-high-performing School)

Pre-test	*Experimental Class*				*Comparison Class*			
	3I (HA) 40 (all in)		3D (LA) 40 (1 absent)		3J (HA) 40 (1 absent)		3H(LA) 40 (4 absent)	
	No.	%	No	%	No.	%	No.	%
6 groups of 4	27	67.5	5	12.8	13	33.3	10	27.8
4 groups of 6	3	7.5	15	38.5	8	20.5	6	16.7
Both 6 groups of 4 and 4 groups of 6	6	15	12	30.8	15	38.5	7	19.4
Incoherent (others)	4	10	7	17.9	3	7.7	13	36.1
Total	40	100	39	100	39	100	36	100
Post-test	*Experimental Class*				*Comparison Class*			
	4I (HA)		4D (LA)		4J (HA)		4H(LA)	
	40 (1 absent)		40 (6 absent)		40 (3 absent)		40 (2 absent)	
	No.	%	No	%	No.	%	No.	%
7 groups of 5	8	20.5	9	26.5	12	32.4	5	13.2
5 groups of 7	9	23.1	7	20.6	2	5.4	7	18.4
Both 7 groups of 5 and 5 groups of 7	21	53.8	16	47.0	23	62.2	24	63.2
Incoherent (others)	1	2.6	2	5.9	0	0	2	5.2
Total	39	100	34	100	37	100	38	100

6.4.4 Results from school-based exams

Students' Primary Two mathematics assessment scores were documented as an entry base for the project. Subsequently, their scores in the semestral examinations, i.e., SA1 taken in May 2004 or M2004 and SA1 taken in Oct. 2004 or F2004 at Primary Three, and SA1 at Primary Four in May 2005 (or M2005) were collected and analyzed. Below are the analyses.

In the high-performing school, the following results were obtained for high-performing classes in the experimental and comparison groups.

1. Regarding students' P2 results, t-test shows that students in the experimental class had significantly lower scores than those in the comparison class.

2. In M2004, t-test shows that students in the experimental class had lower scores than those in the comparison class but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using statistical procedure GLM (general linear model).

3. In F2004, the result shows that students in the experimental class had significantly lower scores than those in the comparison/high class; both classes had similar trend (M2004→F2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

4. In M2005, the result shows that students in the experimental class had significantly lower scores than those in the comparison class; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

Figure 6.1 shows the statistical results in terms of the students' average scores graphically. Note that in the figure, test 1 refers to the end exam of P2 or just P2, while test 2 was used to represent the exam taken in May 2004 or M2004, test 3 for F2004, and test 4 for M2005.

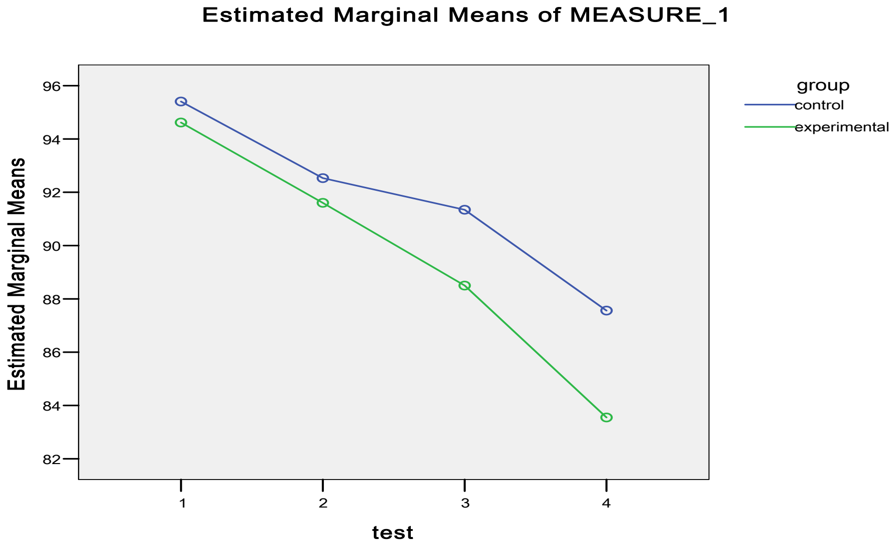


Figure 6.1. School-based exam results for experimental and comparison classes in high-performing school (high ability classes)

For the low ability classes in the experimental and comparison groups in the high-performing school, the following results were obtained.

1. Regarding students' P2 results, the results of t-test reveal that students in the experimental class had higher scores than those in the comparison class, but no significant difference was detected.

2. In M2004, t-test results show that students in the experimental class had higher scores than those in the comparison class but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

3. In F2004, the statistical results show that students in the experimental class had higher scores than those in the comparison class but no significant difference was detected; both classes had similar trend (M2004→F2004: increase in scores) and in terms of the extent of changes, no significant difference between the classes was detected using [GLM].

4. In M2005, the results show that students in the experimental class had lower scores than those in the comparison class but no significant difference was detected; both classes had similar trend (F2004→M2005: drop in scores) but the extent of changes in the experimental class was significantly greater than that in comparison/low class [GLM].

Figure 6.2 presents the statistical results in terms of the students' average scores graphically.

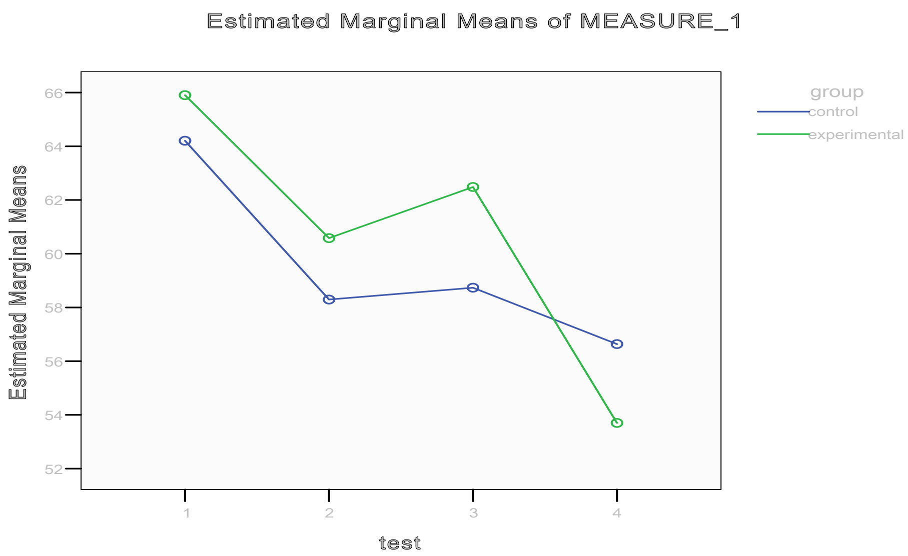


Figure 6.2. School-based exam results for experimental and comparison classes in high-performing school (low ability classes)

Now let us turn to the non-high-performing school. Again, we start with the experimental and comparison classes with high ability students.

1. Regarding students' P2 results, t-test shows that students in the experimental class had significantly lower scores than those in the comparison class.

2. In M2004, t-test shows that students in the experimental class had lower scores than those in the comparison class but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) but the extent of changes in experimental class was considerably smaller than that in the comparison class using [GLM].

3. In F2004, the results reveal that students in the experimental class had lower scores than those in the comparison class but no significant difference was detected; both classes had similar trend (M2004→F2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

4. In M2005, the results show that students in the experimental class had significantly lower scores than those in the comparison class; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

Figure 6.3 depicts the statistical results in terms of the students' average scores graphically.

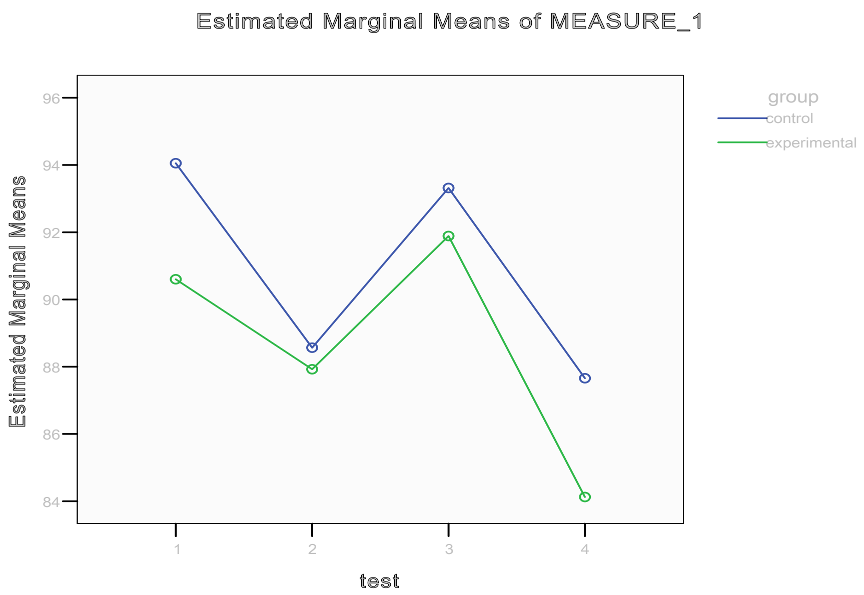


Figure 6.3. School-based exam results for experimental and comparison classes in non-high-performing school (high ability classes)

For the low ability classes in the experimental and comparison groups in the non-high-performing school, the following results were obtained.

1. Regarding students' P2 results, t-test results reveal that students in the experimental class had higher scores than those in the comparison class, but no significant difference was detected.

2. In M2004, the results show that students in the experimental class had higher scores than those in the comparison class but no significant difference was detected; both classes had similar trend (P2→M2004: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

3. In F2004, the results show that students in the experimental class had higher scores than those in the comparison class but no significant difference was detected; both classes had similar trend (M2004→F2004: increase in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

4. In M2005, the results show that students in the experimental class had lower scores than those in the comparison class but no significant difference was detected; both classes had similar trend (F2004→M2005: drop in scores) and in terms of the extent of changes, no significant difference between the two classes was detected using [GLM].

Figure 6.4 shows the statistical results in terms of the students' average scores graphically.

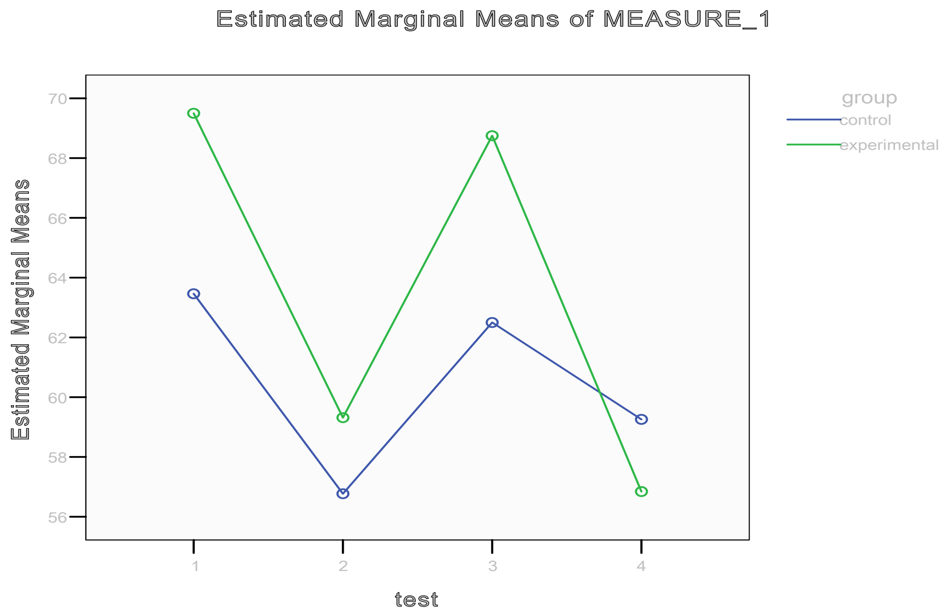


Figure 6.4. School-based exam results for experimental and comparison classes in non-high-performing school (low ability classes)

In summary, the results of students' performances in the experimental groups are mixed in comparison to the comparison classes. Again, as is well known, students' school exam scores are affected by many different factors. Therefore, we think it seems safe to say that the intervention had non-adverse or neutral influences on students' academic achievement as measured in standard school exam results. Further study is needed to obtain a more accurate conclusion in this aspect.

6.5 Conclusions, Implications, and Recommendations

In conclusion, based on the information presented in this chapter, we can conclude the following for each of the research questions.

1) What are the effects of using Journal writing strategies in the mathematics classrooms on students' achievement in both cognitive and affective domains?

The findings of this study have not shown non-adverse or neutral effects on the cognitive achievement of the students. Although the findings on the affective domain are not as positive as one might expect, there are underlying factors such as the change in the teachers for the later part of the research, the way the tasks were enforced could have lent to the students' not-so-positive experience of mathematics learning as well as the use of journal writing in the everyday curriculum context.

Results from analyzing students' writings show that students' ability to perform and compute using multiplication and division do not necessarily mean the ability to explain and interpret concepts. Students seemed to find it easier to show teachers their computational skills than to explain in their own words the concepts or ideas or their thinking. Students' writings also seemed to reflect their rigid thinking about thoughts and views about mathematical procedures. It is also worth noting that teachers felt they had benefited personally in being able to consciously review own teaching and abilities.

2) How can journal writing strategies be effectively integrated in the teaching practice of mathematics teachers?

Based on the results, the researchers had the following main recommendations for the integration of journal writing into the classroom.

First, to improve students' ability to communicate in writing about mathematics journal writing should be emphasized in the normal daily teaching and learning in the classroom. The exposure and the given opportunity for students to express their thinking, ideas and knowledge to teachers and peers will encourage reflection on their own learning and thus develop further depth of understanding.

Second, journal writing has a place in the learning of mathematics in primary schools. As a teacher comments: "Through journal writing, I am able to have a better understanding of my students' weaknesses". Journal writing can help enhance the teaching strategies of the teacher. Journal writing helps students to communicate to teachers that what is taught may not be learned in the same way by all students. It also provides an avenue for teachers to reflect on their teaching strategies as can be seen in the majority responses from their students. Teachers can then ask themselves: "Are my students thinking or are they just regurgitating procedural aspects of mathematics learning?"

Third, if journal writing is to be implemented in the schools, it is imperative that teachers be trained to craft journal prompts. It is also necessary for teachers to be given the support and understanding the tasks thus crafted must be an integral part of the teaching and learning of mathematics.

Finally, journal writing in mathematics cannot and must not be taken and used as an add-on to the existing written tasks. Instructional leaders need to recognise the benefits of journal writing and pave the way for teachers to help them help their students in understanding their thinking and analysis in their learning of both mathematical concepts and problem-solving.

Chapter 7 Results and Findings (V): Performance Tasks (Secondary I)[8]

7.1 Introduction

As mentioned earlier, performance-based assessment is one of four new assessment strategies studied in the MAP project. At the secondary level, two secondary schools, one a high-performing school and the other a non-high-performing school, participated in the study.

This chapter reports the results of a sub-study which focused on the integration of performance tasks into the teaching and learning of mathematics in the high-performing school.

7.2 Research Questions and Conceptual Framework

7.2.1 Research questions

Following the overall framework of the MAP project and focusing on the use of performance assessment tasks, this sub-study is intended to address the following four specific research questions.

1) What are the impacts of using performance assessment tasks in math classroom instruction on students' attitude toward math and math learning?

2) What are the impacts of using performance assessment tasks in math classroom instruction on students' academic achievement in conventional assessment?

3) What are the impacts of using performance assessment tasks in math classroom instruction on students' academic achievement in unconventional assessment?

4) How to use performance assessment tasks effectively in daily math classroom instruction?

It is hoped that the study can provide research-based evidence on the potential influences of using performance assessment tasks on both students' learning and teachers' classroom teaching so as to help teachers better align assessment practice with the desired educational goals and hence improve the quality of teaching and learning.

7.2.2 Conceptual framework

'Performance assessment task' is not a new term in education. Nevertheless, there is no consensus on its definition. According to Buechler (1992), the emergence of performance assessment movement was due to the fairly widespread dissatisfaction with high-stakes multiple-choice tests. Kane, Khattri, Reeve, and Adamson (1997) believed that performance assessment, compared to multiple-choice tests, was more pedagogically valuable and could more accurately reflect students' achievement. Gripps (1994) claimed that, in the states, performance-based assessment was often regarded as any type of evaluation which was not multiple-choice or standardized testing. However, such a definition is rather broad and it almost covers all the alternative assessment modes (e.g., journal writing, project work, etc.).

Recognizing the limitation of traditional assessment modes, the TIMSS (Third International Mathematics and Science Study) team included performance assessment as one important component in their international comparison; they referred it to the use of integrated and practical tasks, which targeted on students' content and procedural knowledge as well as students' ability in using knowledge for reasoning and problem solving (Harmon et al., 1997). The Wisconsin Education Association Council (1996), at the root of the meaning of the word "performance", defined performance assessment as the one requiring students to demonstrate their skills and competencies by performing or producing something. The central idea in Stenmark's (1991) definition about performance assessment is that such an assessment mode shall assess what students actually know and can do. It is clear that while all these researchers tended to differentiate performance assessment from the traditional assessment, they often had had different concerns and focuses, which, to some extent, reveals that there are diverse aspects involved in performance assessment.

To be more applicable to Singapore school education context, this project defines performance assessment tasks as those having two distinguishing characteristics: authentic in context and open-ended in approaches and answers. In fact, these two aspects are to a large degree lacking in traditional assessment, which consequently often receives criticism (e.g., see Howe & Jones, 1998; Wu, 1994).

The authenticity of a problem, according to *Assessment Standards for School Mathematics* (National Council of Teachers of Mathematics [NCTM], 1995), is the degree to which tasks are faithful, comprehensive, and complex, which can be found in important, real-life performances of adults that are non-routine yet meaningful and engaging for students. It is believed that tasks with these features could engage students in applying knowledge and skills they have learned in classrooms to real-world challenges, and help them appreciate the usefulness of mathematics.

The open-endedness in of a problem includes two aspects: (1) multiple venues of access or ways of solutions, and (2) multiple acceptable answers to the problem. It is believed that solving open-ended problems is more challenging than close-ended ones students usually encounter in their school work, and normally requires higher-order thinking.

While the authenticity is more about task context, the open-endedness can be used more to assess students' ability in problem solving. As a result, all the performance assessment tasks used in this study are contextualized in a real-world scenario/story; they can be approached in various ways, from trial and error to more systematic methods, and ended with different answers (not just in different representation formats). Correspondingly, two particular performance rubrics were set to evaluate students' work: *Approaches*, which examine the effectiveness of the strategies used, and *Solutions*, which examine the number of different answers being achieved. Besides the two, *Representation* is also included as a third rubric, which focuses on how the problem solving procedures have been articulated. A sample performance assessment task is shown below:

> The sum of both a mother and her daughter's age was 30 in one particular year. Given that when the sum increased to 40, the mother's age was greater than three times of her daughter's, find all possible age of the mother when she gave birth to her daughter. (Hint: Let x be the mother's age when she gave birth to her daughter, hence when the daughter is y years old, the mother's age is $x + y$ years old)

This task is about age difference between a mother and her daughter, which is about the topic of inequality. The task shows students a fact that although the difference between two persons' ages is fixed, the ratio between the two persons' ages changes every year. It is a common sense but probably ignored by many people. To solve this task, one can use "trial and error" to find some possible ages for the mother without much difficulty. However, solely depending on this method, one would have difficulty in getting all the answers. In fact, there are a total of seven different answers. More systematic methods can be employed, such as making a systematic list or establishing appropriate inequalities. Using the more effective methods, one can solve the task with much less difficulty. Due to the fact that the task can be solved via different methods, students can approach the task at their own levels and solve the task accordingly.

The Singapore mathematics syllabus also emphasized the importance of students' applying mathematics in solving real-life problems and being engaged in open-ended investigations in their learning of mathematics (Ministry of Education [MOE], 2002). However, an analysis of two widely-used Singapore secondary mathematics textbooks revealed that fewer than 2% of textbook tasks were authentic and about 2% were open-ended (Fan & Zhu, 2000). Moreover, according to Schoenfeld (1992), the beliefs that learning of mathematics has little or no relation to the real world and that any mathematics task has one and only one answer are actually very common among students. In this connection, the study has both theoretical and practical significance.

7.3 Methods

As mentioned earlier, this report focused on one participating secondary school, which was identified as a high-performing school, as it was randomly selected from the 50 best performing secondary schools according to year 1999 and year 2002 GCE "O" Level Examination results released by the (MOE). The following sections provide the detailed information about the participants, including both students and their mathematics teachers, the instruments used in the project, as well as the procedures of data collection and data analysis.

7.3.1 Participants

In Singapore, secondary education consists of four types of courses (i.e., Special, Express, Normal Academic, and Normal Technical) catering for students of different learning ability. The high-performing school involved in this study only offered Special and Express courses. Correspondingly, two Secondary One classes (one as experimental and one as comparison) from each course were randomly selected to take part in the project. As a result, the two experimental classes were expected to receive chapter-based interventions on performance assessment tasks during regular mathematics lessons for about three school-terms starting from early 2004, while the corresponding comparison classes were taught as usual during this period of time. No significant difference was found between the two classes within the same stream level in terms of students' Primary School Leaving Examination (PSLE) overall scores (special: t [76] = .81, p = .42; express: t [77] = -.139, p = .89) and mathematics grades (special: U [38, 40] = 747.00, p = .84; express: U [40, 39] = 649.00, p = .12).

Table 7.1 provides the profiles of the four classes of students and their mathematics teachers.

The table shows that the two classes at each stream level were taught by different teachers with basically equivalent professional background during the year 2004. However, due to some unforeseen reasons, starting from January 2005, the teacher teaching the Special/Experimental class had to take over the Special/Comparison class as well. Given the change, the teacher was advised not to use the intervention tasks in the comparison class so as to keep the teaching practices unchanged in both the special classes in terms of interventions. In addition, the teachers from the experimental classes received training and guidance on how to use performance assessment tasks in teaching before and during the intervention from the researchers.

Table 7.1. A profile of participating students and their Mathematics teachers

	Special Classes		Express Classes	
	Experimental	Comparison[1]	Experimental	Comparison
No. of Students	38	40	40	39
Girls	24	25	11	20
Boys	14	15	29	19
	Math teachers			
Gender	Male	Male	Female	Male
Length of teaching experience	3 mth	9 mth	3 mth	3 mth
Qualification	M.Eng, PGDE[2]	MSc	PGDE	PGDE

Note. [1]In year 2005, the teacher teaching the Special/Experimental class also took over the Special/Comparison class. [2]PGDE stands for Postgraduate Diploma in Education.

7.3.2 Instruments

Largely consistent with other components of the MAP project, the following four main instruments were also designed for this sub-study: attitude survey questionnaires, "new strategy task" tests, intervention task worksheets, and interviews with students and teachers.

Attitude survey questionnaires

Two questionnaires were designed, one for the pre-intervention survey and the other for the post-intervention survey, to measure students' attitude toward mathematics and mathematics learning as well as their experience with performance assessment tasks. All the students attempted the pre-survey in February 2004 and the post-survey in May 2005.

Both the questionnaires are comprised of two parts. The first part, identical in both the surveys, consists of 22 items which focuses on students' perceptions about the subject of math and the value of math learning and includes four specific aspects: general view towards math and math learning, anxiety level in math learning, students' perceptions of their own performance in math, and students' beliefs about the usefulness of math. A nine-point scale ranging from "disagree totally" to "agree totally" is employed in this part.

The second part in the pre-intervention questionnaire (6 items) was intended to measure students' experience with various alternative assessment tasks (3 items relevant to performance assessment tasks) in their mathematics learning before intervention with a six-point scale on frequency, while the corresponding part in the post-intervention one (16 items) focused on students' feeling about using performance assessment tasks in math learning using the same scale as the first part.

A pilot study of the pre-survey questionnaire was conducted on 8/9 January 2004, involving 56 secondary one students from two schools from the population but not the sample schools. The questionnaire was improved based on the results of the pilot test. In particular, two items, one on general view and one on belief were finally removed from the pilot version in order to enhance the reliability level of the two sub-scales (general view: from .91 to .92; belief: from .78 to .80), with an average reliability being .85. Furthermore, some items were rephrased so as to become more readable to the students.

"New strategy task" tests

Similar to the questionnaires, two sets of parallel "new strategy task" tests, a pre-test and a post-test, were designed. The use of the pre-test enables researchers to have a better understanding about students' entry levels in math problem solving, whereas the use of the post-test enables researchers to detect possible changes of students' ability in problem solving after three school terms with or without being exposed to performance assessment tasks in mathematics learning. Both tests contain three open-ended tasks, with one being also authentic.

A pilot study of the "new strategy task" pre-test was conducted on 19 February 2004 with 35 secondary one students (Express: 17, Normal Academic: 18) from one school. Based on the students' feedback, necessary modifications on test tasks were made. The modified tasks were again piloted by a group of 36 Normal Academic secondary one students from another school on 1 March 2004. As a result, while about 60% of the students felt that the tasks were challenging to them, all the students had no difficulty in understanding the tasks. Some minor modifications were further made in finalizing the pre-test items.

"New strategy" intervention tasks

As the main intention of the study was to integrate performance assessment tasks into daily classroom teaching and learning, the design of the intervention tasks strictly followed the stipulated school scheme of work (SOW). All the intervention tasks meet both the criteria as described earlier: authentic as well as open-ended. One or two performance assessment task worksheets were designed for each chapter by the researchers. Although the two experimental classes in the study take different types of courses (i.e., Special or Express), they follow the same SOW and use the same textbook. Therefore, the worksheets are identical for both the classes. However, the classroom teachers were encouraged to make necessary modifications about the tasks so as to better fit the students and teaching scheme.

The researchers observed most interventions to monitor how the performance assessment tasks were carried out in the classrooms (some were video taped with the teachers' agreement). The observations were also useful for the researchers to improve the design of future performance assessment tasks.

Interviews with teachers and students
The interviews with teachers and students were conducted mainly for getting information about the participants' experience and understanding about the use of performance assessment tasks as well as their opinions/suggestions on the use of new type of tasks in teaching and learning. While the interview questions for both teachers and students have few differences, those for teachers are more in a teaching perspective and those for students are more in a learning perspective.

Both the teachers from the experimental classes received an individual interview. Six students from each experimental class (two high performing, two average performing, and two low performing) were recommended by the teacher from the respective class to attend the student interviews, in which these students were grouped in three for each session. As a result, two sessions of teacher interviews and four sessions of student interviews were carried out.

7.3.3 Data collection

The pre-intervention questionnaire survey was conducted on 16 February 2004 for all four classes with a response rate of 99.4% and the post-intervention questionnaire survey was on 13 May 2005 for the two Express classes and 18 May 2005 for two Special classes with a response rate of 90.0%.

The pre-test on "new strategy task" was conducted in March/April 2004 with a response rate of 98.7% and the post-test was on May 2005 with a response rate of 89.7%. As the "new strategy task" tests assess more about students' ability in solving unconventional tasks, it is also interesting to investigate how students perform in solving conventional tasks. With the participating teachers' assistance, we were able to collect all the four classes of students' PSLE overall scores and math grades (Exam A), year 2004 school mid-year math exam scores (Exam B), year 2004 school final-year math exam scores (Exam C), as well as year 2005 school first math common test scores (Exam D).

Although both Special and Express students follow the same SOW and use the same math textbooks, there still exists some differences between the two courses for students in terms of their learning ability. Accordingly, the respective math teachers took different paces in teaching. As a result, while the Special/Experimental class implemented 12 interventions during the whole intervention period, only 4 interventions have been conducted in the Express/Experimental class. In most cases, the interventions were recorded with field notes, or audio/video taping. Students' work was collected by the classroom teachers and then handed to the researchers for evaluation. After grading, a copy of students' work with researchers' comments was returned back to the individual student for their revision.

The interviews were arranged once all the above surveys/tests were completed. The interviews with the students from both the experimental classes were conducted on 24 May 2005 and each session (three students from the same class per session) lasted about 30 minutes. The teacher from the Special/Experimental class received the interview also on 24 May 2005 (after school) for about 60 minutes, and the one from the Express/Experimental class had the interviews during school time on 26 May 2005, with each interview lasting about 25 minutes. All the interviews were recorded by audio taping and filed notes.

7.3.4 Data process and analysis

The data from the two survey questionnaires were analyzed using quantitative methods. The descriptive analysis methods (e.g., frequency and percentage) was employed to describe students' overall perceptions about math as a subject and their learning of math when they just entered secondary school and one and half a year after with or without being exposed to math performance assessment tasks. Mann-Whitney U tests were used to examine the possible differences between the experimental classes and their corresponding comparison classes before and after the intervention period so as to enable researchers to detect the impact of using performance assessment tasks on the experimental students' attitudes.

Students' work in the two "new strategy task" tests was graded based on task-specific rubrics by two independent researchers. The inter-reliability was calculated by the Intraclass Correlation Coefficient (ICC) on absolute agreement. As a result, the reliability on three aforementioned performance criteria (i.e., Approaches, Solutions, and Representation) over the three tasks for the two tests ranged from .98 to 1.00, with an average being .99. Similar to the analysis for the questionnaire data, the rubric-based grades from the "new strategy task" tests were analyzed by descriptive statistics to investigate students' overall performance at class levels before and after intervention period. A Mann-Whitney U test was employed to identify possible differences between the experimental and comparison classes in the two tests. A Wilcoxon Signed-Ranks test was used to detect the change in students' grades from the pre- to post-tests. Moreover, possible differences on the changes between the pair of classes at each stream level were examined by a Mann-Whitney U test to identify the potential relationship to the intervention program.

Students' PSLE overall scores and math grades were compared by t-tests and Mann Whitney U tests respectively to ensure the equivalence of the experimental and comparison students in terms of students' academic performance. The followed three school exam scores were analyzed again by t-test to investigate the changes in the differences between the pair of classes when the intervention in the experimental classes was ongoing. A 2×2 ANOVA with time (Exam B vs. Exam D; Exam B vs. Exam C; Exam C vs. Exam D) as a within-subject factor and treatment (experimental vs. comparison) as a between subjects factor was used to detect the potential effects of the intervention program on the students from the experimental classes.

As the interview data were collected mainly in an audio format, all were transcribed. Using qualitative method, the data allow researchers to discover how these teachers and students viewed about the new type of assessment strategy, which is to determine from by questionnaire surveys or achievement tests. It is believed that the evidence from the interviews can triangulate what have been revealed in the above quantitative data so as to strength the findings of the study.

7.3.5 Limitations of the study

To investigate the effects of using performance assessment tasks on students' learning of mathematics, this study involved one experimental class and one parallel comparison class at each stream level. Ideally, the comparison class should not be exposed to performance assessment tasks during the intervention period, whereby the experimental class do. However, in year 2005, the mathematics from the Special/Experimental class took over the corresponding comparison class due to unforeseen reasons. Although the teacher was explicitly asked not to try out intervention tasks in the comparison class, the experience of the teacher working with the experimental class could still influence his teaching in the comparison class one way or another, which can, to a greater or lesser extent, affect the results of the study.

According to the research design, the experimental students should be exposed to the performance assessment tasks in a systematic and scheduled way. Nevertheless, due to some unexpected school activities, it was often very difficult for the teachers to do so in delivering the tasks to the students. As a result, in the first semester, the two experimental classes only managed to carry out one intervention task each. Moreover, performance assessment tasks were not only new to the students but also the classroom teachers; therefore, it is understandable that teachers need time and practice to acclimatize themselves with the new strategies and different lengths of the acclimation periods by different teachers were expected. In terms of the numbers of intervention being carried out (12 vs. 4), the teacher from the Special/Experimental class seems to get used to the new assessment strategy much faster than the one from the Express/Experimental class.

In addition, while this study introduced performance assessment tasks to math teaching in the experimental classes, at the school level those students are still assessed based on the traditional assessment practice for their school performance grading and reporting. Such an inconsistency in the two domains could also have some negative influences on the results of the study (see more discussions in the next section).

7.4 Results and Discussions

The main findings of the study were reported below, based on the four data set described earlier.

7.4.1 Results from questionnaire surveys[9]

Students' general view toward math and math learning

To examine students' general perception about math and their learning of math, six items were designed. The data revealed that students from all the four classes overall provided positive responses to these items in the pre-survey, except one case that the Special/Comparison class had negative view on spending time on math (Q16). In the post-survey, while the students from the two experimental classes continued to give overall positive responses to all the six items, those from the Special/Comparison class expressed an unwillingness to spend time in studying math (Q16) as well as attending math lessons (Q19) and those from the Express/Comparison class believed that math was hard for them (Q5).

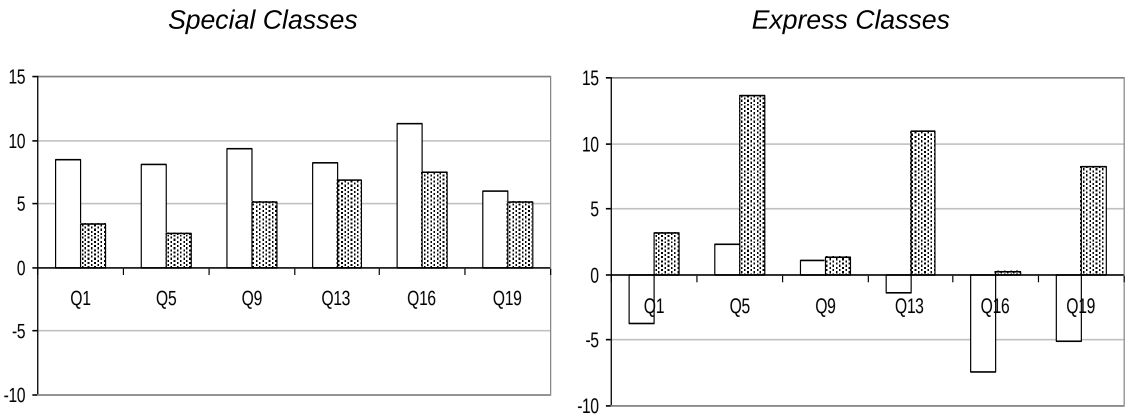


Figure 7.1. Differences in mean rank between experimental and comparison classes by streams about the general views of math and math learning

Note. (1) The blocks without dots represent the data from the pre-survey and those with dots for the post-survey. (2) The difference is calculated by the experimental class' mean rank minus away the comparison class'; the results for the items with negative statements were made necessary reversions before the computation.

Figure 7.1 showed that in the Special stream, the students from the experimental class had more positive views than those from the comparison class in both the pre- and post-survey on all the six items. In the Express stream, the experimental class provided more negative responses than the comparison class in the pre-survey, but the difference became in favor of the experimental class in the post-survey.

From the figure, one can easily conclude that the differences between the two Special classes became smaller from the pre- to post-survey on all items. Moreover, while the Special/Experimental students appeared significantly more willing to spend time on studying math than those from the comparison class in the pre-survey (Q16: U [38, 39] = 522.50, $p < .05$, $r = .25$[10]), the responses between the two classes had no significant differences in the post-survey (Q16: U [33, 29] = 363.50, $p = .101$). A possible reason for this change is that the students from the experimental class had more opportunities to work on performance assessment tasks, which meant much more effort and more time compared to conventional school math tasks. Such a difference might result in some negative impact on students' attitudes toward spending time on math. However, it was not observed in the comparison between the two Express classes. In particular, while the students from the Express/Comparison class showed more willingness to spend time on math than the experimental class in the pre-survey, the two classes of students provided similar responses in the post-survey. The inconsistent results between the two streams could be related to the fact that the Express/Experimental class was not exposed to as many performance assessment tasks as the Special/Experimental class (4 vs. 12) so that the potential impact of using performance assessment tasks on the Express/Experimental class students might not be as evident as that on the Special/Experimental class students, if any. Furthermore, it is found that while the students from the two Express classes had no significantly different responses on all the six items in the pre-survey, those from the experimental class expressed significantly more positive to the hardness of math (Q5: U [39, 38] = 478.00, $p < .01$, $r = .35$) and good feelings about math (Q13: U [39, 38] = 531.00, $p < .05$, $r = .28$) in the post-survey.

Examining the average rating by the students in the two surveys, one may note that the students from all the four classes, in general, appeared more negative in the post- than pre-survey. More specifically, the students felt that learning of math was less enjoyable (Q1) and less interesting (Q9). Furthermore, they were less willing to attend math lessons (Q19). These negative changes were somehow understandable, as when students were promoted to higher grades, math content became more difficult, which needs more effort from students. However, some exceptions were found in the Express/Experimental class. Compared to the responses in the pre-survey, the students from this experimental class provided more positive responses on the items about hardness of math (Q5), good feeling about math (Q13), and willingness to attend math lessons (Q16). The positive changes may be relevant to those students being given opportunities to work on performance assessment tasks in math learning. Nevertheless, to what extent the changes can be attributed to the use of such a new assessment strategy needs more in-depth investigations.

Students' anxiety level in the learning of math

There were six items in the surveys assessing students' anxiety about their math learning. While the students from the two experimental classes gave overall positive responses to all these items in the pre- and post-survey, those from the comparison classes provided negative responses to some items in the post-survey but all positive in the pre-survey. In particular, the comparison class students in the post-survey expressed that they were somehow under a terrible strain in math lessons (Q2) and lack confidence when doing math (Q20).

In all cases but one, the experimental class students had more positive views than their corresponding comparison peers, especially the Special stream, in both the surveys (see Figure 7.2). Moreover, many of such between-classes difference became greater from the pre- to the post-survey.

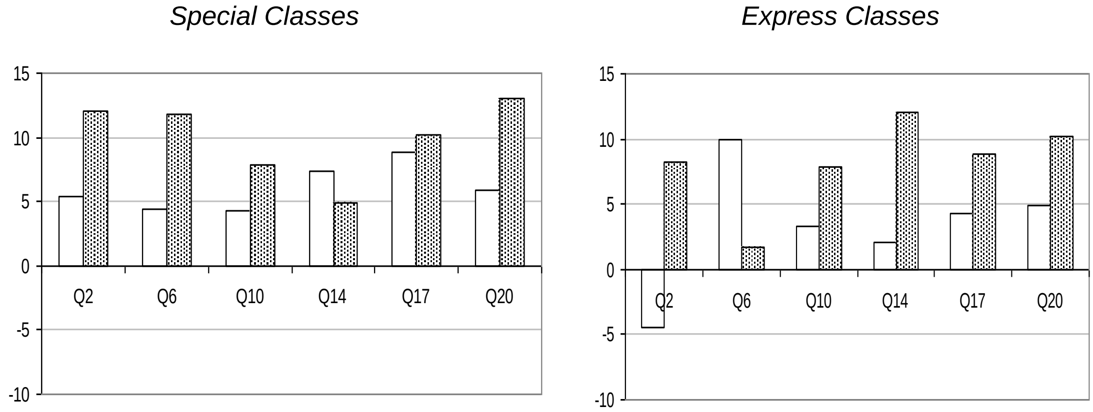


Figure 7.2. Difference in mean rank between experimental and comparison classes by streams in anxiety about math learning

Note. (1) The blocks without dots represent the data from the pre-survey and those with dots for the post-survey. (2) The difference is calculated by the experimental class' mean rank minus away the comparison class'; the results for the items with negative statements were made necessary reversions before the computation.

The post-survey revealed that on four of the six items, the Special/Experimental class students provided significantly more positive responses than their counterparts. More specifically, the Special/Experimental class students were significantly less stressful (Q2), less afraid of (Q6), less nervous (Q17), and more confident about math (Q20) than the corresponding comparison class (see Table 7.2) and their effect sizes range from .33 to .39 with an average being .37.

Table 7.2. Comparison between experimental and comparison classes on anxiety level items by streams

	Special classes		Express classes	
	Pre-survey	Post-survey	Pre-survey	Post-survey
Q2	654.50	293.50^{2}	673.50	583.00
Q6	675.00	297.00^{2}	566.50^{1}	709.00
Q10	677.50	357.00	697.00	589.00
Q14	600.00	403.50	720.50	508.00^{1}
Q17	570.00	322.00^{1}	676.00	570.00
Q20	645.00	278.00^{2}	665.00	545.00^{1}

Note. $^{1}p < .05$, $^{2}p < .01$. The values in the tables are obtained by Mann-Whitney U-test, which examines the differences in the ranked positions of ratings between the experimental and comparison classes.

In general, the students from all the four classes provided more negative responses in the post- than pre-survey. In particular, it was the case on five items for the Special/Experimental class, six for the Special/Comparison class, four for the Express/Experimental class, and five for the Express/Comparison class. Based on the results, one can easily conclude that while the students from all the four classes became more anxious about their math learning from the pre- to post-survey, the negative changes by the students from the experimental classes, especially the Special one, were much smaller than those from the comparison classes. The differences, to some extent, could attribute to the experimental class students' experience with performance assessment tasks, as these tasks are more challenging than conventional school math tasks students and provided students more opportunities to be engaged in higher order thinking, which in turn helps students to establish more self-confidence about math and therefore become less afraid of math.

Students' perceptions of their own performance in math

This category also consists of six items. The data showed that the students from all the four classes in general had positive views about their own performance in math in both the surveys with one exception for the Express/Comparison class in the pre-survey and two for the two comparison classes in the post-survey. In particular, the students from the Special/Comparison class indicated that they did not like solving challenging math problems in the post-survey (Q21) and those from the Express/Comparison class felt that they were not good at math in the two surveys (Q11).

It is found that from the pre- to post-survey, students became more negative toward their own math performance and it is common across all the four classes. The negative movement could be related to the fact that more challenges and stresses came along when students moved to higher grades. However, as shown in Figure 7.3, the students from the Special/Experimental class were significantly more willing to attempt challenging math tasks (Q21) than their counterparts in the post-survey (U [33, 29] = 333.00, $p < .05$, $r = .30$) but it was not the case in the pre-survey (U [38, 40] = 671.00, $p = .369$). It appeared that those students' experience with performance assessment tasks had some positive influence on students' self-confidence about their own performance.

In general, the students from the experimental classes provided more positive responses than those from the comparison classes, although the differences became smaller from the pre- to post-surveys in most cases. In particular, the Special/Experimental class students had significantly stronger belief that they could do well in math (Q15) than their counterparts in the pre-survey (U [38, 39] = 546.50, $p < .05$, $r = .26$), but it was not the case for the post-survey (U [33, 29] = 398.50, $p = .252$). Such a change seems understandable as the students from the Special/Experimental class tried out quite a number of math performance assessment tasks and they often cannot be easily solved. Such experience might lead those students to less believe that they could do well in math in the post- than pre-survey. Similar changes also appeared in the comparison of the results from the two Express classes, where the experimental class students expressed significantly higher confidence in their math ability (Q11) in the pre-survey (U [39, 39] = 487.00, $p < .01$, $r = .36$) but not in the post-survey (U [39, 38] = 574.50, $p = .083$).

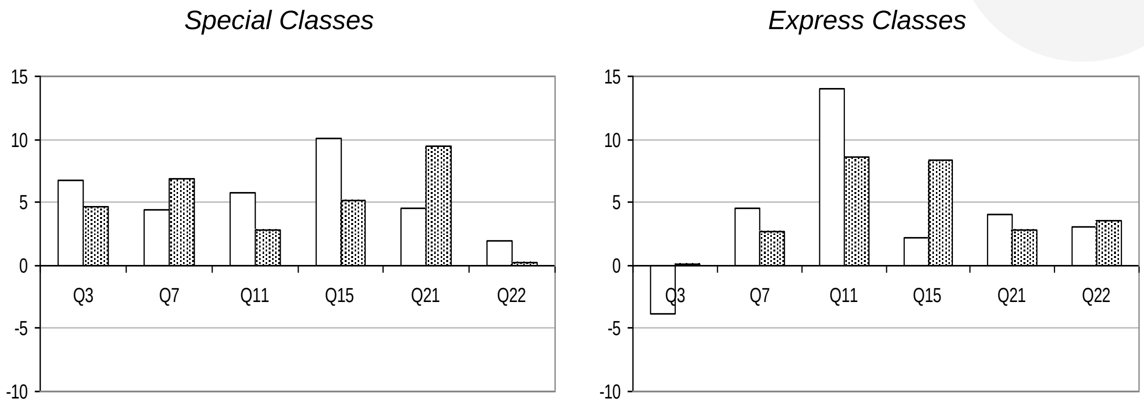


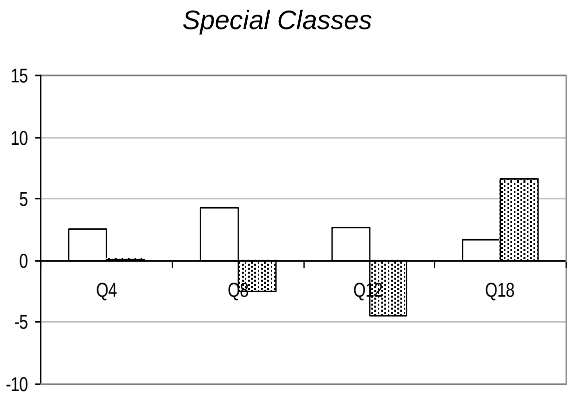
Figure 7.3. Difference in mean rank between experimental and comparison classes by streams in perception about their own math performance

Note. (1) The blocks without dots represent the data from the pre-survey and those with dots for the post-survey. (2) The difference is calculated by the experimental class' mean rank minus away the comparison class'; the results for the items with negative statements were made necessary reversions before the computation.

Students' belief about the usefulness of math

Four items were set to assess students' belief about the usefulness of math. In both surveys, the students from all the four classes provided positive responses on all the items. However, it is also found that compared to the responses in the pre-survey, those in the post-survey were generally more negative, except two cases in the Express/Experimental class. In particular, these students agreed more with the statement on the usefulness of math (Q4) and the importance of math (Q8) in the post- than pre-survey. The overall negative changes might be related to the fact that when students moved to higher grades, math became more abstract and appeared further away from students' daily life, although fortunately the students from all the four classes still gave a generally positive view to the usefulness of math in the post-survey. Moreover, the nature of performance assessment tasks contextualized in the real world seems also attribute to the results that none of the Special/Experimental class students disagreed with the usefulness of math (Q4) in the post-survey, while about 5.3% of them did so in the pre-survey.

Inconsistent with the results from the other three aspects reported earlier, while the Special/Experimental class students presented more positive responses on all the four items than their counterparts in the pre-survey, more negative views were given in the post-survey (see Figure 7.4). In particular, the comparison class students agreed more with the importance of knowing math nowadays (Q8) and the meaningfulness of studying math (Q12) than those from the experimental class. Such an unexpected result might be related to the fact that while the experimental class students were given many opportunities to work on performance assessment tasks which involved real life application of math knowledge, the skills they learnt from such tasks were virtually never assessed in their school examination. The inconsistent practice could bring students to downgrade the value of studying math and in turn believe that studying math (which was never tested) was somehow wasting time.



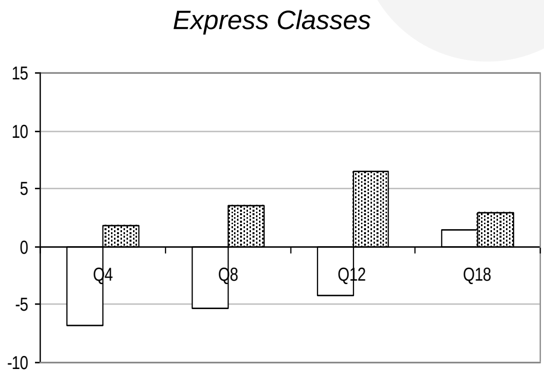


Figure 7.4. Difference in mean rank between experimental and comparison classes by streams in belief of the usefulness of math

Note. (1) The blocks without dots represent the data from the pre-survey and those with dots for the post-survey. (2) The difference is calculated by the experimental class' mean rank minus away the comparison class'; the results for the items with negative statements were made necessary reversions before the computation.

In contrast, at the Express stream, the comparison between the two classes was in favor of the comparison class in the pre-survey but the experimental class in the post-survey. One might argue why the negative influence did not find in the Express/Experimental class. As mentioned earlier, given the fact that the students from this class had only tried a small number of performance assessment tasks, the potential impact of using the new strategy might not be detected.

Students' experience with real-life and/or open-ended tasks
The questionnaire also set a number of items assessing students' experience with tasks having particular features, including tasks in real-life context and tasks with multiple answers or approaches. All the students from the four classes were requested to report the frequency of doing such tasks in the pre-survey. The two comparison classes were again asked to answer these items in the post-survey so as to allow researchers to examine the consistency between the actual classroom practice and research design, where the two comparison classes were expected to be kept as intact in terms of using performance assessment tasks.

The results from the pre-survey showed that the students from all the four classes in general had overall similar experience in doing the tasks with the aforementioned features. They worked on these types of tasks either in a monthly base or a weekly base. The comparison between the pre- to the post-survey, in terms of percentage changes, indicated that the students from the two comparison classes did not have many changes in these experiences.

Students' perceptions about performance assessment task and its usage
Due to the fact that only the students from the two experimental classes received the intervention on using performance assessment tasks in math learning, only these two classes of students were asked to report their views about performance assessment task and its usage. A total of 16 items are designed for the purpose, consisting of three aspects – perceptions of the authenticity and open-endedness of performance assessment tasks, viewpoints about the challenge of the new strategy, as well as beliefs of the usefulness of doing performance assessment tasks.

Authenticity and open-endedness are the two most important characteristics of performance assessment tasks which distinguish them from other assessment modes. Six items are related to this aspect. The data showed that the majority of the students provided either neutral or positive responses to all the six items, which indicated that the students overall welcomed the specific features of performance assessment tasks, including multiple approaches of the tasks (Q26) and the authenticity of the task contexts (Q30 & Q31). Furthermore, the students believed that doing performance assessment tasks helped them to be more creative (Q27) and systematic (Q32). However, it seems that the students were still uncomfortable with the open-endedness in final answers. It is understandable that with previous school experience, students were often only requested to provide one and only one correct answer to each task and they had already been used to such tradition and felt comfortable about it. The open-endedness in final answers, in contrast, brought students not only challenges but also somehow confusions, as commented by one student in the interview (more detailed in later section).

Regarding the difficulty of performance tasks, it appears that the students generally felt that doing performance assessment tasks were challenging. In particular, more than 65% of the students claimed that they had to think harder in doing performance task (Q28) and 63% believed that it was time-consuming (Q35). About 30% of the students felt lost when doing performance assessment tasks (Q29) and 47% asked for hints' help.

There are four items asking for students' views on the benefit of using performance assessment tasks for their math learning. The results showed that the majority of the students did not have negative views toward the usefulness of working on performance assessment tasks. In particular, the students believed that doing performance assessment tasks helped them in learning math (Q25) and it was not wasting of their time (Q38). However, about 49% of the students disagreed that doing performance assessment tasks could help them learn math better (Q36) and only 38% indicated that they would like to have more performance assessment tasks for their math lessons (Q37). It could be due to the fact that performance assessment tasks have not been included in the school assessment system so that the students were unable to "see" the immediate benefit of doing such tasks and at least it does not seem to help them to get higher marks in the conventional school tests. Consequently, the students might devalue the usefulness of doing performance assessment tasks in their math learning and some became unwilling to have more of such tasks in future study.

7.4.2 Results on students' math academic achievement

This study evaluated students' math academic achievement in two ways. One is students' performance in conventional tests, which were represented by students' school-based examination results (e.g., PSLE math grades and Semester Assessment [SA]). The other is students' performance in the pre- and post-tests on "new strategy task", which were designed by the researcher.

Results on school-based exam scores

As reported earlier, there was no significant between the two Special classes as well as the Express ones in terms of students' PSLE overall scores as well as math grades (Exam A). It indicates that the two classes at the respective stream level were about equivalent for their math academic achievement. The analysis of the three followed test scores were to examine students' learning progress over the intervention period, which are 2004 mid-year exam results (Exam B), 2004 final-year exam results (Exam C), and year 2005 first common test results (Exam C). The analysis revealed some significant changes, shown as Figure 7.5.

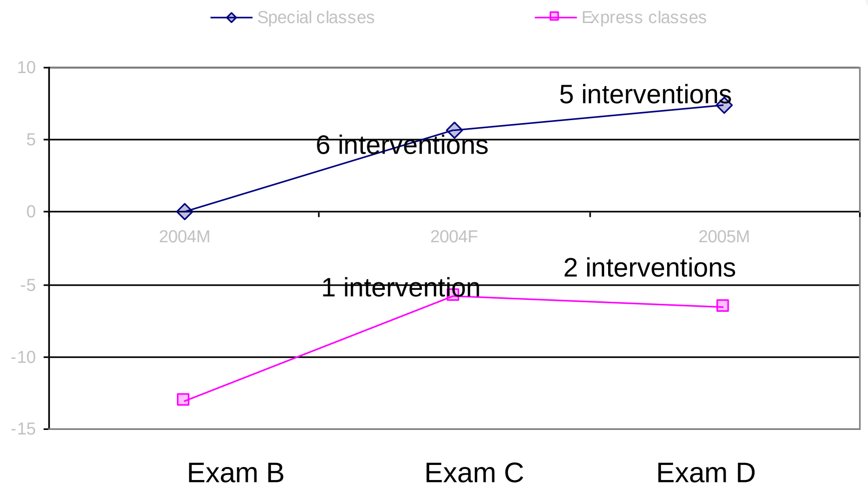


Figure 7.5. Difference between experimental and comparison classes on school-based exams

Note. (1) The difference is calculated by experimental class' mean score minus its comparison class'. (2) The numbers on the lines indicate the no. of performance assessment tasks (interventions) being carried out between the particular time points.

From Figure 7.5, we can see that the change for the Special/Experimental class compared to the comparison class was in a more positive way. As reported earlier, there was no significant difference between the two Special classes in students' PSLE math grades. The comparison at Exam B again showed the equivalence between the two classes (t [75] = 0.015, p = .988). As a matter of fact, before Exam B, the experimental class only managed to carry out one intervention and no great change for the experimental class was expected. More interventions were carried out after Exam B, as indicated in the diagram. The comparison of students' performance at Exam C and Exam D revealed that the differences between the two Special classes became considerably large (Exam C: t [74] = 1.935, p = .057, r = .22; Exam D: t [74] = 1.961, p = .054, r = .22). A repeated measurement analysis of variance between Exam B and Exam D revealed that there was a significant interaction between time and treatment effects (F [1, 74] = 8.392, p < .005, r = .32) and the effect size is about medium, which is in favor of the experimental class. Furthermore, the analysis showed that the significant interaction actually occurred between the period from Exam B to Exam C (F [1, 73] = 6.682, p < .005, r = .29) whose effect size is also nearly medium.

Inconsistent with the results from the comparison of students' PSLE math grades, the Express/Comparison class students performed significantly better than those from the experimental class in Exam B (t [72] = -6.802, p < .001, r = .62). It indicates that the two classes may not be equivalent in terms of students' academic performance. In fact, in the follow-up two school exams, it is consistently observed that the comparison class students had significantly better performance than the experimental class students (Exam C: t [76] = -2.355, p < .05, r =.26; Exam D: t [77] = -2.141, p < .05, r = .24). However, one can easily find that the differences between the two classes became smaller gradually. A repeated measurement analysis of variance between Exam B and Exam D was also conducted for the two Express classes. As a result, a significant interaction between time and treatment effects (F [1, 76] = 6.533, p < .05, r = .28) was detected in favor of the experimental class and the effect size is about medium. A further analysis showed that the interaction occurred in the period from Exam B to Exam C (F [1, 76] = 13.077, p < .001, r = .38). However, similar to the Special/Experimental class, the Express/Experimental class was also only able to implement one intervention before Exam B. Therefore, it seems hard to claim that the experimental class students' experience with performance tasks was the sole factor for such a result. Fortunately, the progress maintained in Exam D, when a total of four interventions had been carried out.

Results from the "new strategy task" tests
Compared to the school-based exams, the new strategy test items are more similar to the intervention tasks. Each test consists of three items with all being open-ended in approaches and answers and one being authentic in context. Moreover, the pre- and post-tests are designed in a parallel way so as to enable the researchers to identify the possible relationship between students' changes in performance and their experience with performance assessment tasks. The students from all the four classes took the pre-test in April 2004 before the intervention started and the post-test in May 2005 when the intervention ended.

In terms of the overall scores in the tests, the results showed that the students from all the four classes made improvement from the pre- to post-test (see Table 7.2).

Table 7.2. Students' mean scores in the pre- and post- new strategy task tests by classes

Classes		Pre-Test	Post-Test
Special	Experimental	17.79	21.00
	Comparison	15.66	21.66
Express	Experimental	18.50	18.62
	Comparison	16.90	20.14

Note. In both tests, full marks are 36.

A Wilcoxon Signed Ranks Test revealed that the changes for all but the Express/Experimental class reached significant level in the favor of the post-test (Special/Experimental: $Z = 3.598$, $p < .001$, $r = .62$; Special/Comparison: $Z = 3.882$, $p < .001$, $r = .75$; Express/Experimental: $Z = 0.455$, $p = .649$; Express/Comparison: $Z = 2.934$, $p < .01$, $r = .51$). However, the Mann-Whitney U-tests did not find any significant between-class difference at the respective stream in either test.

As all the tasks are open-ended in nature, it would be more meaningful to exam students' performance in terms of their usage of effective strategies, the number of answers obtained, as well as the representation of solutions than their overall scores. It is believed that the analysis on these sub-domains could provide more in-depth information on how students approach and solve such challenging tasks, especially those from the experimental classes. A brief description of the three performance rubrics is listed in Table 7.3.

Table 7.3 A brief description of general rubrics by approaches, solutions, and representation

	Level 0	Level 1	Level 2	Level 3	Level 4
Approaches (decision or strategy about approaching tasks)	• No attempt or • No evidence of a strategy	• Strategy is ineffective and could not lead to any correct answer	• Strategy could lead to correct answer but not systematic (e.g., guess and check)	• Strategy shows partial systematic pattern	• Strategy is effective that would lead to a complete set of answers
Solutions (no. of answers obtained)	• No correct answer obtained	• Only one correct answer obtained	• More than one correct answer obtained	• At least 50% of the full answers obtained	• A complete set of answers obtained
Representation (documentation of problem solving procedures)	• No attempt or • Working is irrelevant	• Working is not clear and hard to read	• Working is not organized so that the approach is not observable	• Working is organized and approach is partially observable	• Working is well organized and approach is fully observable

Regarding the approaches employed by the students, the data revealed that in most cases the students were able to use more systematic/effective methods in the post- than pre-test. That is, more students received a mean score over 2 on this performance scale in the post- than pre-test (Special/Experimental: 79.4% vs. 42.1%; Special/Comparison: 82.8% vs. 23.7%; Express/Experimental: 66.7% vs. 45.0%; Express/Comparison: 74.3% vs. 28.2%). Such an improvement reached a significant level for all but the Express/Experimental class (Special/Experimental: $Z = 3.929$, $p < .001$, $r = .67$; Special/Comparison: $Z = 4.115$, $p < .001$, $r = .79$; Express/Experimental: $Z = 1.756$, $p = .79$; Express/Comparison: $Z = 3.562$, $p < .001$, $r = .62$). Moreover, the progress made by the Express/Comparison class was significantly greater than that by the corresponding experimental class ($Z = 1.999$, $p < .05$, $r = .24$), which was not the case for the two Special classes. However, no significant between-class difference was found at the respective stream level in terms of using effective strategies in both the tests.

As stated earlier, all the tasks in the tests were open-ended; that is, each task contains more than one correct answer, as listed in Table 7.4.

Table 7.4. No. of correct answers to test items

	Pre-test	Post-test
Task 1	7	10
Task 2	56	25
Task 3	2	2

Due to the fact that the last task in each test has only two answers, a task-specific rubric on solutions was set for the particular tasks, shown in Table 7.5.

Table 7.5. A task-specific rubric for tasks in the pre- and post-test

Task 3	Level 0	Level 1	Level 2	Level 3	Level 4
Solutions (no. of answers obtained)	No correct answer obtained	Only partial correct answer obtained, i.e., getting correct central number(s). Or Answers obtained just by switching surrounding numbers without changing the central numbers	One complete answer with different central number obtained	Two correct central numbers with one complete answer obtained	Two complete answers with different central numbers obtained

The analysis revealed that compared to the pre-test, the percentages of students who stopped at obtaining one correct answer were much smaller in the post-test for all but the Express/Experimental class (average score ≤ 1.33), shown in Figure 7.6. In fact, those three classes made significant improvement in getting multiple answers from the pre- to post-test (Special/Experimental: $Z = 3.39$, $p < .001$, $r = .58$; Special/Comparison: $Z = 3.787$, $p < .001$, $r = .73$; Express/Comparison: $Z = 2.639$, $p < .01$, $r = .50$). However, the between-class comparison on students' changes on the "solutions" scale from the pre- to the post-test at the respective stream level did not display any significant difference. In addition, similar to the results on the "approaches" scale, in both the pre- and post-tests, no significant between-class difference was detected on the "solutions" scale at each stream level.

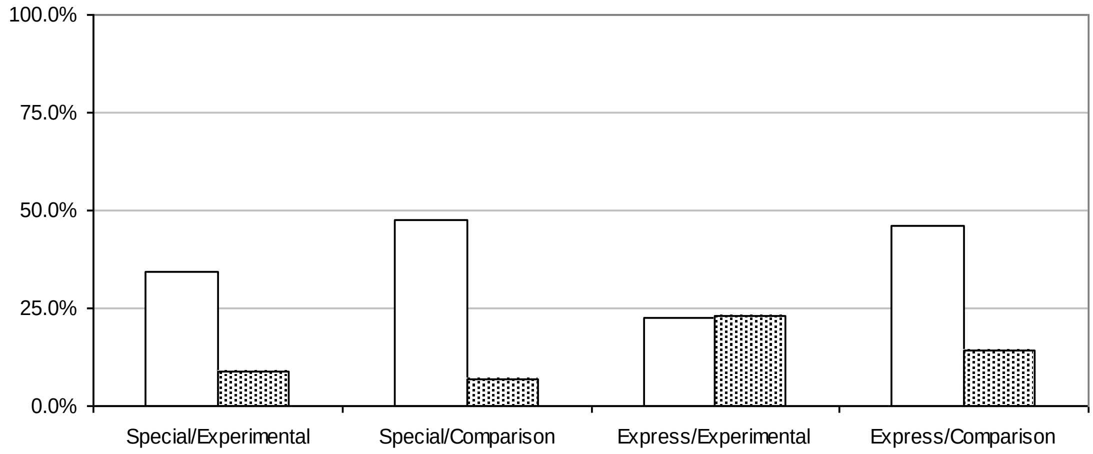


Figure 7.6. Percentages of students whose mean scores on "solutions" were no more than 1.33 in both the tests

Note. The blocks without dots represent the data from the pre-test and those with dots for the post-test.

It is believed that presentation is also an important skill in problem solving. Therefore, although it is not a focus of the intervention program, how students represent their solutions in the new strategy tests was also examined. The results revealed that the students generally did not have significant changes in their representation from the pre- to post-test except that the Special/Comparison class seemed to have significantly better performance in the post-test ($Z = 2.720$, $p < .01$, $r = .52$). However, such progress made by this comparison class was not significantly greater than that by its experimental class, whereas significant difference was revealed between the two Express classes in favor of the comparison class (U [39, 33] = 440.500, $p < .05$, $r = .32$). On the other hand, consistent with the results from the other two performance scales, no significant difference was found between the experimental classes and their comparison classes in either the test on the “representation” scale.

The results showed that all but the Express/Experimental class made great improvement from the pre- to post-test, especially in the aspects of using effective strategies and getting multiple answers. On the other hand, one may also notice that while there was no significant between-class difference within the respective stream in both the tests, the advantage was on the side of the two experimental classes in the pre-test but the two comparison classes in the post-test. One possible reason for such an undesirable result is that the students from the experimental classes knew well about the project and they were clear that all their grades on performance assessment tasks would not be counted into their school records. Correspondingly, those students may not treat the post-test as seriously as their peers from the comparison classes who were just given the test without further information. However, the fact that the pair of classes at the same streams did not show significantly different performance in the post-test somehow indicated that using performance assessment tasks surely did not harm students’ learning of math.

7.4.3 Results from interviews

The math teacher and six students (two high performing [H], two middle performing [M], and two low performing [L]) from each of the experimental classes received interviews in late May 2005. Different from the teachers who received individual interviews, the students from the same class were grouped in three containing all the three performance levels to be interviewed. As a result, six interview sessions were conducted.

Recognition of the specific features of performance assessment tasks
As described earlier, the performance assessment tasks used in this research project had two specific characteristics which made them different from the regular math tasks. One is that the tasks must be open-ended in both approaches and answers, while the other is that the tasks are in a real world context. As the fact that the two teachers attended a training seminar before the intervention started, they should be aware of the nature of the performance assessment tasks. In contrast, the students were exposed to the tasks but not informed explicitly about the particular features. Therefore, one interview question to the students is targeted on students’ awareness of the specific features of the performance tasks. From the students’ view, the most distinguishing feature of the performance assessment tasks is the open-endedness in approach and answers. Some relevant comments are listed as follows:

- There are many different ways to answer the solutions and then there are also many different answers, so we can actually try to use different ways and actually try to find the solutions. (S_H1)
- You need not have to stick to one solution to certain question. (S_L1)
- More answers, because normal time you get either get it correct or you get it wrong. … this one is more answers. Like more than one answer, more explanations, it’s like we do not have a lot of word problems you know, our normal … We can think in more way and we know that actually a question can have more than one answer. We are exposed to more … we are exposed to different types of maths. (S_M2)
- There are many methods and there are many solutions to these questions (E_H1)
- Many answers as well … word problems (in textbook), they only had one answer … but this has lot of answers (E_L1)

However, for this feature, the 12 students had different views. Two students (S_H2, S_M2) claimed that some of those questions were not relevant to their normal math. One student (S_L2) expressed his dislike about the multi-answers and the reason he pointed out is “because you have found one of the answers, then if we check with other people for the answers, we thought that either one of us was wrong”. In fact, the open-endedness in approaches also brought some confusion, as S_M2 commented. Some students appreciated the open-endedness of the performance assessment tasks. For example, E_M2 believed that doing such questions made them more flexible. S_L1 felt that it gave them a different view of math; say, math is not just doing homework questions. Similarly, S_H2 stated that it brought them more experience and they were more aware of such questions now. E_H1 stated that doing the performance tasks made him more creative and think widely.

Regarding the open-endedness feature, the two math teachers generally believed that it benefited their students' math learning. However, they very much concerned about their students' ability. The teacher from the Special/Experimental class claimed that such tasks would be more beneficial for *strong class*, who has potential in them by their own, than those who need specific instructions. The other teacher from the Express/Experimental class commented that it was helpful for students to have the concept that one question could have more than one solution. Nevertheless, she found that many of her students were not capable to cope with the performance assessment tasks yet as they could not even grasp the basic fundamental concepts. In her view, the performance assessment tasks compared to other normal math tasks involved more higher-order thinking skills, which might be appropriate only for about 20% of the students in her class to try out.

The real world setting is another important characteristic of the performance assessment tasks. It is found that not many students recognized this particular feature, although they generally noticed that there was difference between the performance assessment tasks and the regular math tasks they encountered in textbooks or school tests:

- These questions require more common sense. (S_M1)
- These questions help to relate maths with the real life, like Painting a Room, let us see how maths actually can help us in real life. (S_H1)
- It is different from out textbooks as in the pictures and situations in which you don't get in your textbooks. Like in your textbooks, there are a mostly examples and formulas and all the questions that come out. But here you see that you use examples of geometry in another way like you can use it in maths. (S_L1)
- (the content is more difficult) than normal class work … like, need to read the question carefully. (E_L1)

Both the teachers were supportive to the real world context. One teacher commented that "it is a different thing from the normal dry concepts pounding worksheets. So in that sense students may … find it interesting". However, sometimes he also found that students did not show interest at all. From his point of view, it was very much gender-oriented. For example, girls may find "supermarket" exercise interesting, while boys' interests may be in soccer. To him, getting students' attention is important as it could ensure students to be highly engaged in the tasks. Therefore, this teacher was very careful about the intro and the exit in delivering worksheets.

The other teacher felt that it was good for students to know that in real life or in practical situations they can have a variety of ways, means, and answers to what they are going to solve. Moreover, she suggested that it might be better if we could start off with simpler tasks but not the higher order type.

Difficulties with using performance assessment tasks

Due to the fact that performance assessment tasks are new to teachers as well as to students, it is understandable that both sides encountered difficulties/challenges when they used the tasks, especially at the beginning. During the interviews, the students and teachers indicated several difficulties they have encountered so far:

1. Difficulty in understanding the questions (from students)
   - Sometimes I don't even understand what the question is asking for … It's just that it is very confusing. But then after a while, (my teacher) explains to us, then we get clear (S_H1)
   - Sometimes I don't really get what the question is trying to ask (S_L1)
   - Some of them do have difficulty in understanding the question (E_T)

2. Difficulty in finding a workable approach (from students)
   - (in normal work), we do linearly lah, like straightforward, … this one to involve more abstract, more complicated (E_H1)
   - you have to gather a lot of information and use it and sometimes like quite lost and we don't (know) where to put what in the question (S_L1)
   - which method has the fastest way of solving it, or which combination of the method let us be like more comfortable and not so tedious, so different methods have different ways of using it, and it will make the problem like, some will make more easy and some will more tedious (S_L2)
   - (difficult) to choose which method to use (S_H2)

3. Difficulty in finding all answers (from students)
   - Not capable to find the answers, we can't finish the things (E_H1)
   - Like one question like think so much … and if you cannot, you do a lot of work, you still cannot solve, you are very irritating (E_L2)
   - It will be frustrating if you cannot get the answer after trying so many methods (E_H2)

4. Time-consuming (from students and teachers)
   - Normally we take very little time, but this one, we need to spend lot of time on these questions, I mean we need to think of other ways (E_H1)
   - It takes too much one (a) single question (E_M1)
   - Just felt time-consuming (E_M2)
   - The reason (for not marking) is the schedule for me with school is already very busy and really don't have time, because it is much more time-consuming than the normal ones, you know and it takes a lot of time as well, as I think it will take a lot of time for the understanding the rubrics and etc … I think the main thing is the time consuming (E_T)
   - They are rather time consuming … (weaker students) do not have a good grasp of the remembered concepts. Due to time constrain, they won't have moved along. When they attempt such worksheets right, then they will be very confused … (as) they already have the basic problem of remembering … (S_T)

5. Difficulty in delivering (from teachers)
   - I think the way to approach, a difficulty is how to bring, how to approach in a simpler manner such that students actually can understand what is going on and how to like, you know, tell them, you know that one that can be many solutions to thing, because I think lots of students now still have this concept that every question must have one solution … so when it comes to like you know you try to tell them there are more, try to think more, you really have to, you know, promote as in provoke their answers to come up. Ya … but my students, some of them like they are very lazy to think … (E_T)

6. Others (from the teachers)
   - the thing on your feet is … the major challenge, then but I think most other teachers, in their normal teaching, will also have to do it basically. But this one the chance of you thinking on your feet is higher because you are probably to type on something that is rather … non-mathematical to go into something

mathematical. So that the thing is there, I think it's a major challenge if I want to mention. (S_T)

- Other challenges, erm, the task of trying to integrate this completely into the curriculum is not easy. (S_T)
- (the other challenge is) non of these are tested in the academic tests, in the academic assessment. So erm, the kids also sometimes question whether do they need to spend so much time on these. They have to place this as second thing loh. … the first question they ask and I cannot lie, is this counted in CA1? The answer is "no", straight away the gear will shift from gear one to gear four … so cannot be helped. … If you tell them that this is somehow let's say 20%, wow, I tell you, you will get very good results. … because they do spend sometime on this and if it is not tested, they felt the time spent is not worth it, or every particular from this selfish point of view. (S_T)

All these challenged both the teachers and students in their use of performance assessment tasks for math teaching and learning. However, after a period of struggling, many of them felt that they were improving in some aspects. For instance, one student from the Special class found that the performance assessment tasks were not that challenging anymore nowadays and he was wondering whether it was because the thing had been simplified or they were getting better (S_L2). Two students from the Express class also felt that they were now able to use different strategies and became more flexible for their thinking, whereas in the past they simply stick to one method (E_M2, E_L2).

As for the teachers, both reported that they became more comfortable in using the tasks in their daily teaching. In particular, the Special class teacher stated that he now at least knew better when and where to slot the tasks in the teaching in order to make his students feel comfortable. Nevertheless, both the teachers also mentioned that they did struggle with the tasks at the beginning of using them and they believed that for a new teacher, one year trial was necessary for the accommodation.

Reception of performance assessment tasks

In general, the interview indicated that both the teachers and students welcome the use of performance assessment tasks. They believed that use of performance assessment tasks could benefit students' learning of math. Regarding the feasibility of replacing certain amount of regular math tasks with performance assessment tasks, all of the interviewees had no objection and felt it won't affect students' performance relevant to cutting down drill and practice:

- This is part of the syllabus, I think it's okay (to replace) because the syllabus is changing anyway … the things that we are doing now in our textbook you can like practice in the exercise book, but these performance tasks, you can't really find them anyway. You know, like you can't just practice this exercise book with this performance task. (S_M2)
- (For replacement), give you one fifth (S_L1) … A quarter (S_H1)
- I think if you want to replace, I think it must at least a 70% or more … otherwise your this will never fly … (if only) 30%, there will be people who do it and people who don't do it, then the thing … will die off natural death in 1 year. … If 70% is too much, I think at least 50%. You must have an overwhelming change in management of this, otherwise it is very difficult. (S_T)
- Maybe (the MOE) can reduce by about one quarter… I felt that if you actually free out the time, you have more time and space for the students to think. … I'm going to have the extra like 25% of time coming out, you can have more time to allow students okay we can really just leave these two periods for you to go through and think through and come about with creative answers for the math solutions rather than, you know, really drill, drill and drill practice things … (E_T)

When come to the issue about a proper frequency of using performance assessment tasks in their daily math teaching and instruction, many of the interviewees suggested the pace of one or two tasks per chapters is workable. Although some believed that performance assessment tasks could be also used for introducing a new topic, most of them regarded the tasks as a kind of enrichment. Therefore, to them, it is more appropriate to use performance assessment tasks after basic concepts have been covered, if not at the end of a chapter. Moreover, some students commented that before trying performance assessment tasks, it is important to go over all important concepts and methods in the relevant chapter.

As performance assessment tasks are one mode of alternative assessment, it is interesting to know whether students accept the idea of including their grades on performance assessment tasks into their CA/SA records. The results showed that all the 12 students did not mind about such a practice with an acceptable range from 10% to 20% of CA/SA total scores. The Special class teacher commented that "*if …the students deemed it as important, right? Should you must let this occupies 15%. This constitutes to the whole thing; 15% of the CA I think, I think 15%, some overlapping, at least half, or else, in any part of time, students have this thinking that this is not part of actually curriculum work, because deep down they still think not so. It's very hard to enforce on this one.*"

Suggestions and comments on use of performance assessment tasks
Reported below are feedbacks we obtained from the interviews about the use of performance assessment tasks

A. Voice from students

Preparation:
1. We go and read up on the subject that you are doing. Example like geometry, you go and read up the textbook and … later when we do it we won't be like so stuck because like say it involves a certain formula you didn't learn about it, then you don't know how to do the question. (S_L1)
2. Maybe like practice 1 hour a day. Just to warm up. … Keep our brains fresh. (S_L2)
3. Really understand the whole chapter first … go through details and then allow them to attempt this (E_H1)
4. Maybe recap on the basic method we know (E_H2)

Strategies:
5. Not to only stick on one solution, can think of other ways. Then the question will be easier to solve and not only one way to do all these questions. (S_L1)
6. Think more. Use more of your brains. (S_M2)
7. Not to stop at only one answer (S_H2)
8. Read the question carefully (E_L1)
9. Be more flexible, try out different ways, … don't give up (E_L2)
10. Think more, … let them apply then, then combination, it will help … Be patient. (E_M2)
11. Think out of the box (E_H1)
12. Think of the methods learnt before (E_H2)

Format:
13. (For group discussion) so that you know one solution and your friend knows another solution … compare the solutions and then use the more efficient one. (S_H1) For example, if you get stuck, you won't get stuck there. You can seek other opinions. (S_L1) Because if we get the same solution, then we know it's the correct answer. …

sometimes there are some places where I don't know how to carry on and then I will ask them (S_M1)

14. Individual you do already and share with the rest of your friends in your group. That's when you do yourself first then later you like share with each other. (S_H2)
15. Group work. … Share our ideas … putting out ideas together (E_M2) … We can open mind, learn from each other (E_L2) We can mix ideas together, maybe we can form more thoughts. (E_H2)
16. Individual, because we have different opinions, she insists on the answer and we insist on the answer. If there are five persons in a group, five of us think of different answers, totally different answers, like that very confusing. (S_L2)

Suggestions for implementation of performance tasks (If I were the teacher):

17. Make the lesson a little bit more interesting, right, they will understand better. (S_L2)
18. I probably will give them an example first. … Like for example like if the equation, it depends on the question really, like the same question but then this change, here a bit, here and there and then like remember the question is like 3 + 5 and then we just teach you how to do. like 2 + 3, and then we just tell them right. It is something like that. (S_M2)
19. Try to make them voice out their solution and share with their class so that their class will get a brief idea what is going on. … Voice out their opinions on the solutions. (S_H2)
20. Give them hints, one hint by one hint, let them find the way to solve the questions (E_L2)
21. Go through the answers … like explain … how … arrive these answers, because not all the students they understand what is written on the answer sheet (E_H2)

B. Voice from teachers

Preparation:

22. First thing, I would think that must let the teacher feel comfortable. Must give plenty of leeway because nobody will know what will happen, when you put something different into the spot. So, teachers should be encouraged to try their best. What I'm saying MOE first thing they no need to send teachers for courses, I don't think so. … it comes secondary. First thing must give the teachers assurance first. They are free to try whatever they need to and as long as they can justify why, they should have the feeling. This is very important. Number two, … this is non-traditional, non-conventional form of assessment, then the conventional performance figures cannot be used to measure this lah. So must also let the teachers know the thing. (S_T)
23. I think, of course, there must be a training (for) teachers … maybe a seminar or a course … to a kind of to tell them … how to teach in that, because I think many teachers are still constrained … need to retain them, … if this is really to be implemented, I think the first thing is to really give a course, I suppose, a course or a seminar to prepare teachers, like how to teach this in a more interesting and creative manner and I think understanding the rubrics at first is very important also, because some teachers maybe like … lost as to how to implement this, how to mark. … (preparation) with more of mentally. (E_T)

Resource:

24. Resources … I'm sure those teachers have their schools' … I mean I went to a few schools, they also have roughly the same thing, also have the CD-Rom, interactive this and that. Plenty of math puzzles here and there. That one is not lacking. What lacking is carrying this thing out. (S_T)
25. I think examples, … maybe like … a benchmark, some samples, teachers have a look first, maybe … a handbook will be useful, handbook on how to go about doing this kind of tasks, how to set, these kind of things. (E_T)

Format:

26. I think it will be better for like students to do it by group actually … I think through discussions, many new ideas … if you could put kids … in a small group, I think maybe through their discussion … they can come out with the answers, because if they individually do it, and they cannot solve it, then they will just give up, where else if they have their peers and things like that to … encourage to thinking, it will be better. (E_T)
27. My class can do independent because it really cuts down a lot of time and I enjoy it because of the flexibility. … Independent is easier, but it's only after a certain number of interventions or a number of performance-wise stage, the classroom expectation, then you can start doing these. … (In) group work, … the girls will discuss with girls, guys will discuss with the guys, never … yeah. (S_T)

Support needed:

28. I would say that if it could help me along … if … superiors are more supportive, and not with a conflicting thing with, then definitely no problem. But sometimes, due to the time taken, then the kids need to come back for extra remedial. It will sort of tell them that this is actually not important and I'm wasting time on it, especially for the weaker ones. They really feel that this is not really important and it won't be tested anyway and they only have trouble testing, I mean learning what is important. Those which they think must be tested and their normal CAs and their academic assessments. So that part is …although (HOD) supports us but of course to him, the other performances are equally important. (S_T)

As these teachers and students were first attempting performance assessment tasks, the difficulties and frustrations they faced could be common to anyone who first tries out the tasks. Therefore, the suggestions and comments from them about how to effectively use performance assessment tasks in math teaching and learning would be valuable not only for research but also school practice.

7.5 Conclusions, Implications, and Recommendations

In conclusion, from the information presented in this chapter, we can conclude the following for each of the research questions.

1) What are the impacts of using performance assessment tasks in math classroom instruction on students' attitude toward math and math learning?

The impact of using performance assessment tasks on students' attitude toward math and math learning is a gradual procedure. Consistent with many other researchers' (e.g., Macnab & Payne, 2003; Wong, Lam, Wong, Leung, & Mok, 2001) findings, the students in this study in general become more negative toward math and math learning in the post- than pre-survey in all the four sub-domains. On the other hand, the study also revealed that the changes were generally in favor of the students from the experimental classes at both the stream levels, especially the anxiety level about math. However, it appears undesirable that the changes on the view about the usefulness of mathematics were preferable to the students from the comparison classes, while contextualization in real life was one important characteristic of performance assessment tasks. Moreover, in the interviews, many students appreciated the particular feature of the new assessment strategies. One possible reason for the seemingly contradictive results is that the students may find that what they experienced in the performance assessment tasks seldom appeared in their regular school math tests; therefore, to them, the normal school math was more important, as it would really be tested. Compared to the performance assessment tasks, the students from the experimental classes may have a stronger feeling that the math they encountered in the regular school learning was farther from their daily life.

2) What are the impacts of using performance assessment tasks in math classroom instruction on students' academic achievement in conventional assessment?

Regarding cognitive domains, the project looked at students' performance in both the conventional assessment (school exams) and unconventional assessment (i.e., new strategy tests). The results showed that the changes in students' performance across three continuous school semester tests (i.e., Exam B, Exam C, and Exam D) were significantly preferable to the experimental classes at both the stream levels. Moreover, the favorite changes actually occurred after the intervention program had been implemented about one school year, when the Special/Experimental class completed 7 interventions and the Express/Experimental class had 2 interventions, and maintained till the intervention ended. Although such a positive result could not solely attribute to students' experience with the new assessment strategy, it seemed clear that the students from the experimental classes did benefit from being exposed to performance assessment tasks.

3) What are the impacts of using performance assessment tasks in math classroom instruction on students' academic achievement in unconventional assessment?

As to the unconventional tests, the students from both the experimental and comparison classes performed better in the post- than pre-tests, not only in overall scores but also in specific performance domains, including use of effective strategies and obtaining multiple solutions. In terms of the progress from the pre- to post-test, there were generally no significant differences between the experimental and its comparison class at the respective stream level, with only one exception being that students from the Express/Comparison class had significantly better improvement on the representation scale in the post- than pre-test compared to their peers in the experimental class.

It appeared that working on performance assessment tasks did not give the students from the experimental classes the advantage in unconventional tests, where all the tasks were open-ended. As discussed earlier, one possible reason for such a result is that the students from the experimental classes well knew that their performance in such tests would not be counted into their school records. Therefore, these students might not treat the post-test as serious as those from the comparison class. In addition, due to the fact that the intervention implemented in this project is only about one and half school year (12 times in the Special class and 4 times in the Express class), whether the relevant impact could immediately emerge need more evidence to verify. Nevertheless, it is clear that there were no negative effect of using the new assessment strategy appeared on students' performance.

In addition, both the surveys and interviews showed that the students welcomed the new assessment strategy. In particular, the students were happy with the nature of multiple approaches as well as the real-life real life task context. However, some students were not comfortable with the open-endedness of final answers, as they were confused and lacked confidence in their answers. The participating teachers, in the interviews, also confirmed that some of their students were not ready to work with the new assessment strategy, although they believed that the experience with performance assessment tasks eventually would benefit their students' math learning. Due to the fact that the project was only intended to introduce but not officially launch performance assessment into daily school math teaching, the new strategy was not included in the school assessment system. It is believed that such an inconsistent practice may be a reason for some students could not see the usefulness of using new assessment strategy for their math learning.

4) How to use performance assessment tasks effectively in daily math classroom instruction?

From the teachers' points of view, the integration of performance assessment tasks into daily math teaching is a promising practice and also a challenge to them. Not only because it is new to them, but also performance assessment tasks, compared to regular school math, involve more higher-order thinking skills and some students who were relatively weak in math may have difficult to take the challenges. Therefore, more preparation work from teachers is requested for both the students and teachers themselves. Furthermore, the teachers commented that external assistance for them to be used to the new strategy, which includes academic guidance as well as affective support, is important to them.

The results from the study suggested that teachers and students were capable of handling performance assessment tasks. As expected the effect of using the new strategy in some cases was not obvious, it is clear that no negative impact was observed. On the other hand, the study also observed some positive effects on students' academic achievement and anxiety about math learning. It is believed that it is a good practice to provide students opportunity to work on problems contextualized in real life and open-ended investigations, which is actually part of the primary aim of student math learning. Nevertheless, further study is needed to investigate whether the aforementioned unobvious impact will emerge in a later time and for how long the positive effects could last. In addition, this study also revealed that the inconsistent practice between the experimental assessment and traditional assessment resulted in some undesirable consequences.

With appropriate help, the two participating teachers could successfully integrate performance assessment tasks into their regular math teaching and they also believed that other teachers could do so as well. After a certain period, the teachers also became comfortable with the new strategy, although different teachers took different length of time to acclimatize themselves. As a matter of fact, the teacher from the Special/Experimental class got used to the tasks faster than the one from the Express class; at the end of the program, the first teacher managed to implement a total of 12 interventions compared to 4 by the second one. It is believed that the integration of performance assessment tasks into regular math teaching does not only help teachers to improve their teaching, but also provide teachers a venue to have better understanding of their students' learning, such as flexibility in using various problem solving strategies and awareness of the likelihood of a task having more than one correct answer, which is hard, if not impossible, to get via traditional assessment modes. Due to the fact that the new strategy indicated here is about assessment, providing prompt and informative feedback to students is important and essential for the effective use of performance assessment tasks. In addition, a corresponding evaluation system should accompany the implementation of the new strategy.

Finally, we would like to point out that this study is just an initial step to investigate the possible effects of using performance assessment on both students' math learning and teachers' teaching. More research on the impact of using such a new strategy on teaching and learning in various aspects is necessary. Moreover, how to effectively and efficiently use the strategy for instruction also needs further exploration and try-out. It is believed that this is a long term process and shall involve constant evaluation and adjustment.

Chapter 8 Results and Findings (VI): Performance Tasks (Secondary II)[11]

8.1 Introduction

As pointed out earlier, the MAP project aims to examine the effect of integrating four new assessment strategies, namely, journal writing, project work, performance tasks and self-assessment in the daily teaching and learning of mathematics.

This chapter reports a sub-study which focused on the effects of using performance tasks (more specifically open-ended and/or authentic tasks) as an assessment strategy in the mathematics classroom in a neighbourhood school.

8.2 Research Questions and Concepts

The specific research questions for this sub-study are as follows:

(a) What are the effects of integrating performance tasks (authentic and/or open-ended) in classroom instruction on students' mathematical achievement in school-based examinations?

(b) What are the effects of integrating performance tasks (authentic and/or open-ended) in classroom instruction on the students' affects towards the learning of mathematics?

(c) What are the effects of integrating performance tasks (authentic and/or open-ended) in classroom instruction on students' problem solving abilities for open-ended questions?

(d) What influence does the use of performance tasks (authentic and/or open-ended) have on teachers and students in the daily teaching and learning of mathematics?

In general, a performance task is often defined as a problem, a project, or an investigation which may be completed as individual, paired, or group work. Students are required to formulate plans, use problem-solving strategies which are crucial in the development of higher order thinking skills, and provide justification as well as explanation for their answers (also see Van de Walle, 2004).

As indicated earlier, the MAP project focused on two specific types of performance tasks, namely, authentic tasks and/or open-ended tasks. In other words, the authenticity and open-endedness are the most important features of the performance tasks used in this study.

For a discussion about the concepts of authenticity and open-endedness, one can refer to Chapter 7. In particular, 'Authenticity' in this sub-study refers to representations of the contexts found in important, real-life situations which is similar to the definition by NCTM (1995).

The authentic open-ended performance tasks employed in this research are integrated into classroom teaching and homework, providing teachers with another means of collecting information regarding the mathematical understanding of their students. By engaging students in the performance tasks, students learn to construct responses to challenging real-life problems. Hence teachers and students alike would have to shift their paradigms in their focus, from product to process.

8.3 Methods

8.3.1 Participants

Participating School
As is well known, in Singapore, all secondary students are placed in four courses, i.e., Special, Express, Normal Academic or Normal Technical courses in the secondary schools at the end of their primary education, based on their Primary School Leaving Examination (PSLE) scores.

The Special and Express course share the same framework, as it forms the basis of mathematics learning and teaching for all local students in the schools, while the mathematics curriculum for the Normal Academic course is designed as a subset of the Mathematics syllabus for the Special / Express courses. (Ministry of Education [MOE], 2000b). The syllabus content for the Normal Technical course is comparable to that of the Normal Academic course with differentiation in the sequence of topics and rigour in the mathematics content.

The school where the research was carried out is a relatively new neighbourbood school with its first year of operation in 2000. Four classes of Secondary One pupils from the school, approximately 160 students, and 4 Mathematics teachers were involved in the research. Two classes served as the experimental group, one from the Express (EE) and the other from the Normal Academic stream (NE), with the corresponding classes (EC and NC) as comparison groups.

As mentioned earlier, MAP selected participating classes at the Secondary One level as it is deemed the turning point for most pupils as they progressed from primary to secondary education. Streaming is a common practice in Singapore after the second year in a secondary school based on their academic results. Students are differentiated and placed in classes with different subject combinations to match their abilities and interest. As the research spans approximately eighteen months, introducing the new assessment strategy at Secondary One would be most appropriate to avoid the later complication of streaming at Secondary Two.

Profile of Students
In the experimental groups, the students have a mean PSLE score of 218 in the Express class and a mean of 179 in the Normal Academic class. The distribution of the mathematics grades of the four classes at the PSLE is shown in Table 8.1.

Table 8.1. The distribution of students' Mathematics grades in PSLE

Course	No. of students	Experimental /Comparison group	Range of PSLE score	Mean	A*	A	B	C	D	E
Express	40	Experimental (EE)	200 - 238	221	12	20	8	0	0	0
	40	Comparison (EC)	204 -222	214.6	0	20	17	3	0	0
Normal	40	Experimental (NE)	173 - 186	179	0	1	5	18	16	0
	39	Comparison (NC)	173 -187	179	0	0	11	16	10	1

There were minimal changes to the experimental and comparison classes at the end of year when students progressed from Sec.1 to Sec. 2, with the exception of one or two lateral transfer cases.

Profile of Teachers
The participating teachers are all trained teachers with postgraduate qualifications after completing their degrees programs. There were no changes of teachers throughout the eighteen months when the study was carried out.

Table 8.2 shows the general information about the participating teachers in all the class. In the table, Teacher EE represents the teacher who was teaching class EE, i.e., the experimental class in the express course. Similarly, Teachers EC, NE, and NC represent these who were teaching class EC, NE, and NC respectively.

Table 8.2. Profile of participating teachers in the study

1. Express course			2. Normal course		
Teacher	Qualification	Length of Teaching	Teacher	Qualification	Length of Teaching
Teacher EE	PGDE, BA	3 years	Teacher NE	PGDE, B.Eng (Civil)	3 months
Teacher EC	PGDE, BA	8 years	Teacher NC	PGDE, BA	2 years 6 months

8.3.2 Instruments and data collection

Consistent to other components of the MAP project, six instruments were designed and/or used for data collection in this study. They are as follows:

1. School semestral examination results for mathematics,
2. Pre- and Post- Questionnaire Surveys,
3. Pre- and Post- tests,
4. “New strategy” interventions,
5. Interviews, and
6. Video and audio taping of lessons when interventions were carried out in the classroom during mathematics instructional time.

Semestral Examination Results for Mathematics
The Semestral Examination mathematics results, Mid-Year 2004 (M2004), End-of-Year 2004 (F2004) and Mid-Year 2005 (M2005) were collected from all the four classes, the experimental and the comparison classes. The data collected were analysed using quantitative methods to measure the effects of integrating the performance tasks on academic results.

Pre- and Post-Questionnaire Surveys
Two questionnaire surveys: pre- and post-surveys, were designed by the mathematics assessment project team. The questionnaires consisted of 2 parts: A and B, and were piloted, refined and administered on the students. The pre-survey was administered on 12 February 2004 before the first intervention was carried out, while the post-survey was administered on 20 May 2005. Both questionnaires were in the Liket-type scale format with a nine-point scale ranging from “disagree totally” to “agree totally”. The main purpose of the two questionnaires was to investigate the students’ attitude towards mathematics and their learning of the subject, as well as their experience with the new assessment strategy, performance tasks prior to and after the implementation of the interventions. The pre-questionnaire survey serves as a baseline for making comparisons in the attitudes of students after the interventions were carried out.

1. Part A of the Pre- and Post- Questionnaire Survey

Part A of the pre- and post- questionnaires is the same and focuses on how the students perceive mathematics and the value they attached to the learning of the subject. The survey can be broadly divided into four categories:

- the students’ general perception towards mathematics and the learning of the subject,
- the level of anxiety experienced in the learning of mathematics,
- the students’ perceptions of their own mathematical competence,
- their beliefs about the usefulness of mathematics.

A nine-point scale, ranging from “disagree totally” to “agree totally” was used so that small changes in the students’ attitudes can be detected.

2. Part B of the Pre- and Post- Questionnaire Survey
Part B of the questionnaire survey consists of a total of nine questions for the pre- and sixteen questions for the post. The objective of Part B in the pre-questionnaire survey is to measure the students’ prior experience with performance task in the learning of mathematics. However, in the post questionnaire, the intent of Part B is to find out the students’ opinions on performance task after the interventions. As mentioned in Chapter 2, a copy of the pre- and post- questionnaire can be found in the Appendices 2.2 and 2.3.

Pre- and Post-Test
Similar to the questionnaire survey, the purpose of having the pre-test is to provide the baseline for measuring changes in the students’ performance on the new types of tasks. At the same time, the researcher can understand the students’ entry level and design subsequent interventions that cater to their needs.

The pre-test, which consisted of three open ended questions, *Cube Arrangement*, *Movie Selection* and *Filling Numbers*, was conducted on 9 March 2004. As participating students had just progressed from the primary to the secondary education when the research study was conducted, the questions only included topics on arithmetic and geometry, the two main topics in the Singapore Primary Mathematics curriculum.

To assess students’ performance in working on the open-ended tasks, task-specific rubrics were designed for each task by the researchers.

The rubric contains three different criteria:

- Mathematics approach/procedure (A): Decisions on how to approach the problem and includes the strategies, skills, and concepts used to solve the task;
- Problem solution(S): The answer(s) to the question(s);
- Presentation (P): Documentation/Explanation of how the problem was solved. This criterion includes the mathematics language and representation used in presenting the solution

Within each criterion, there are five levels of attainment with descriptors explaining the expectation of performance for each level.

The pre-test was first piloted on 19 February 2004 with another school. Eighteen Secondary One Express students and seventeen Secondary One Normal Academic students participated in the pilot test. During the pilot study, each student was assigned only two of the three open-ended problems found in the new strategies task test. Eighteen students were assigned with the task on *Cube Arrangement* and *Movie Selection* and seventeen students with the task on *Filling Numbers* and *Movie Selection*. In this pilot study, students were able to complete the two tasks within 20 minutes but had expressed difficulty with the comprehension of some terms, such as the word, ‘proposal’ in the task. A 10-minute informal interview with the students after the test revealed that most students did not have any prior experience in solving such open-ended tasks before but they expressed interest in attempting more of such tasks.

With the feedback from the pilot test, modifications to the three open-ended tasks were made. For example, the term ‘proposal’ was changed to ‘combinations of movies’ to ensure greater clarity in the questions. On 1 March 2004, a second pilot study was carried out with thirty-six Normal Academic students from another school. For the second pilot test, students were requested to work on all the three refined open-ended tasks within the timeframe of twenty minutes. The outcome of the second pilot showed that students had little difficulty comprehending the tasks but reflected that the duration of twenty minutes was insufficient to complete all the tasks. Based on that feedback for the second pilot study, it was decided that the duration of thirty minutes would be a reasonable timeframe for students to complete all the three tasks. Besides, some students had also expressed that they had spent too much time working on the first question, *Filling Numbers* thereby they did not have sufficient time to work on the other two. Hence in the final copy of the pre-test, the sequence of the questions is as follows order: *Cube Arrangement*, *Movie Selection*, and *Filling Numbers* as indicated in the Appendix 8.1.

To enhance the inter-rater reliability, two researchers of the MAP team first marked 10 random selected students’ answer scripts individually, based on the pre-designed task-specific rubrics. The average Intraclass Correlation Coefficient, ICC, was 0.839. A fair agreement was reached after the researchers held discussions about the rubrics and made further refinements followed by another round of independent marking. The average ICC increased to 0.994.

The post-test, which consisted of three similar questions to the pre-test, Cuboid Arrangement, Song Selection and Filling Numbers, was conducted at end of May 2005. The students’ performance was assessed based on the same three criteria with a 5-level scale: their approach and reasoning, solution and presentation (which include mathematics language and representation). No pilot test was carried out before the implementation of the post-test as it was deemed to be parallel to the pre-test.

“New strategy” intervention

The main intention of this study is to integrate performance tasks into the teachers’ daily mathematics teaching. The authentic and/or open-ended performance tasks are designed to align with the mathematics syllabus and the weekly scheme of work.

As the two experimental classes, EE (Express) and NE (Normal Academic), in the school were of different streams and the performance tasks were prepared with the expectation that the respective mathematics teachers would make the necessary modifications to cater to the varying ability and needs of the students in both streams. Draft designs of the performance tasks aligned with the subsequent mathematics topics were sent to the corresponding teachers via email who had a choice of administering the performance task to their class depending on the appropriateness of the task and their schedule.

Classroom observations and videotaping during the interventions were made with the consensus of the teachers. All the authentic and/or open-ended performance tasks which the teachers carried out in this research study were crafted by the researchers. All performance tasks consist of two parts, A and B. Working on the basis that student might not be familiar with solving open-ended problems or the context of the problem, part A was namely a warming up exercise. The questions in part A were simple and served to familiarize students with the context of the problem as well as to review the basic concepts and skills involved before exposing them to the authentic/open-ended performance task of part B. For the Express class, EE, ten performance tasks which covered a wide range of content area were attempted in this research study.

Task 1: 'River Cruise' was carried out on 21 April 2004
The performance task, River Cruise, was conceptualised based on the rates charged for the river cruise at Boat Quay. It was specifically designed to align with the topic on algebraic expression which was covered in the Secondary One syllabus. The four sub-topics included in this performance task are: (1) using letters to represent numbers, (2) expressing basic arithmetical processes algebraically, (3) substituting numbers for letters in formulae and expressions, and (4) manipulating simple algebraic expressions. The scenario in this open-ended problem is about a school teacher who needs to bring a class of pupils for a cruise, optimising on the expenditure as well as the number of boats for the cruise is the main focus of the task. With the different variables, the students were supposed to come up with a plan and make a decision, justifying their choice.

Task 2: 'What is the Time now?' was attempted on 25 May 2004
The task required students to demonstrate their ability to formulate an algebraic expression by exploring the angle formed between the hour and minute hands of the clock as they progressed with time. The first part consisted of a warming up exercise where questions were scaffold to lead students to formulate the expression when both the hour and minute hands of the clock formed an angle of 180˚. Students were provided with the manipulative, a clock to help them derive the two different algebraic expression; before and after 6.00pm. In the follow up question of the performance task, students were required to find the time when both hands of the clock overlapped.

Task 3: 'Painting a Room' was completed on 12 July 2004
The task '*Painting a Room*' provides students with the opportunity to relate the mathematical concepts on area and perimeter, to the real –life context of painting the walls and waxing the floor, based on given dimensions. Besides the arithmetic computation, the cost involved and the number of coats of paint and wax, the task required the skill of drawing a balance, written explanations and diagrams when coming up with a quotation as a painter.

Task 4: 'Physical Anthropology' was conducted on 11 August 2004
The performance task was crafted to align with the topic on ratio and proportion. Students were required to find the ratio of the humerus (length of the long bone in the arm that runs from the shoulder to the elbow) to the height of a person. The extension to the task would be finding the best physical part of the body to estimate the height of a person, which is characterised by the arm span. This performance task is very meaningful as it draws the attention of the students to the fact that the performance of a sports athlete at an elite level is affected by his physical makeup.

Task 5: 'Time Deposit' was carried out on 18 August 2004
This performance task requires students not only to use their mathematics skills to calculate the interest accrued over a period of two years, but also to choose the best scheme that maximises the returns.

Task 6: 'National Flag' was conducted on 30 September 2004
This performance task drew the students' attention to the national flags of various countries: Israel, Australia, Malaysia and Singapore. A common symbol that is found on the flags of different countries is the star. The Australian flag contains a seven point star which is very different from the fourteen point symbolic star that appeared on the Malaysian flag. Besides tapping on their conceptual understanding and mathematical skills on the sum of interior angles of polygon as enclosed by the 'stars', this task also exposed students to the national flags of various countries.

Task 7: 'Singapore Currency' was carried out on 19 January 2005
The context of this task relates the physical attributes of the local currency, notes and coins with the physical size of Singapore. In this task, students were required to calculate

the number of $5 currency notes that is required to cover the entire land area of Singapore. In working through the task, students will be able to develop a deeper understanding of the standard form and the rationale for learning the topic.

Task 8: 'Size does matter' was carried out on 8 April 2005
To illustrate the concept of scale, students were treated to a short introductory paragraph on how more than six thousand characters from Sun Tzu, The Art of War, were creatively engraved on the pole-hilt of ancient miniature weapons carved out of ivory. A parallel situation is drawn where students are required to use their mathematical knowledge to creatively rescale the map of Singapore to fit onto badges of different geometrical shapes.

Task 9: 'Building Bricks' was conducted on 21 April 2005
The underlying mathematical concept embedded in this task, 'Building Bricks', is the solving of quadratic equations.

Task 10: 'Decorating your Kitchen' was attempted on 15 May 2005
The concept on Factorisation and Expansion was succinctly built into this task. There are five different conditions when considering this quadratic expression: $p^2+pq+36$. Students were encouraged to systematically explore the possible values of p and q.

For the Normal Academic class, NE, three performance tasks were attempted in the study. Two performance tasks were the same as the Express class, with no modifications made, and only one varied.

Task 1: 'Shuttle Run' was carried out on 18 May 2004
The task was designed to be integrated with the lessons on Rate and Speed. The performance task, 'Shuttle Run' sits well with this subtopic as it is one of the requisites in the National Annual Fitness Award (NAFA) test that all secondary students would have to take to determine their fitness level.

Task 2: 'River Cruise' was carried out on 4 August 2004
Same as the Express stream.

Task 3: 'Singapore Currency' was carried out on 3 February 2005
Same as the Express stream.

For illustration purpose, Performance Task 9 for the Express students, its solutions, and the assessment rubrics are provided in Appendices 8.2, 8.3, and 8.4 respectively. Readers who are interested may contact the researchers for all the other performance tasks and their assessment rubrics that were carried out in the study.

The implementation of "New Strategy" Interventions
Ideally, the performance tasks should be seamlessly integrated into classroom teaching and learning, minimizing interruptions in instructional time and maximizing the impact on students' learning. Hence, all tasks administered by the mathematics teachers of classes EE and NE, during the lesson were designed to align with the scheme of work.

Students were initially requested to read the questions and attempt the warming up questions individually, based on their own understanding. Group discussion follows where students compare their answers with one another. The teachers then spent a proportion of their instructional time on whole-class discussion or question-and-answer sessions where at the end of that discussion, students volunteered their solutions or were called upon to answer part of the performance task by explaining and showing their workings on the whiteboard. The workings presented by their peers, then became a focal point for further discussion where students evaluated the processes and the approaches adopted. Other strategies or procedures which varied from those presented were also discussed in the class.

Students were then requested to complete their performance tasks individually. Those who were unable to complete the entire performance task within a two period timeframe would finish it as homework. Answer scripts were collected on subsequent days as instructed by their mathematics teachers.

The researchers would separately assess the answer scripts of the performance task based on a task specific rubric. Before finalising the different levels of attainment for the three criteria, the researchers would meet briefly to discuss on answers which deviate from the proposed solution produced by the researchers or points excluded from the task specific rubric and finalise the scores.

For each student in the EE and NE classes, a customized print-out of the following items was prepared:

- the levels of attainment,
- comments that were reflected on the answer scripts by the researchers,
- the different criteria with brief descriptors and levels of attainment that are found in the rubric and
- the proposed solution

The print-out was given to the teachers to be returned to the students within a timeframe of two weeks from the date of submission of answer scripts for most tasks.

Interviews
At the end of the interventions, interviews were conducted with the two teachers and students from the experimental classes, EE and NE. Six students from each experimental class:- two high, two average and two low ability students were interviewed using a structured set of questions prepared by the MAP project team which individual researchers could adapt whenever necessary. The objective of the interviews was to gather further information on the students' perceptions of authentic and/or performance tasks in addition to the data collected from the pre- and post- questionnaire, pre- and post- test and the answer scripts of performance tasks. The students were interviewed in groups of threes to provide a more supportive environment and each session lasted approximately half-an-hour.

Similarly, the teachers involved in the experimental class, EE and NE, were also interviewed using a set of structured interview questions which researchers in the MAP project can adapt for use. The interviews were conducted to obtain information on their perceptions of performance tasks, the limitations in the implementation process and recommendations regarding the use of performance tasks within the Singapore system. All interviews were audio recorded and transcribed. As described in Chapter 2, the structured interview questions can be found in Appendices 2.4 and 2.5.

Video and audiotapes
All interventions carried out in the classroom were video and audio taped with the consensus of the teachers involved. By video and audio taping, the research hoped to triangulate information gathered from the various sources of data: the questionnaire survey, interviews, pre- and post- test. However, the analysis of the lesson structure carried out in the class would not be included in this dissertation.

8.4 Results

Below we report the results and findings obtained from this sub-study to address the four research questions as mentioned at the beginning of this chapter.

8.4.1 Results on students' academic achievement

To address the first research question on effects of the interventions on students' academic achievement in the school-based examinations, the mathematics results for the three semestral examinations, Mid-Year Examinations 2004 (M2004), Final-Year Examinations 2004 (F2004) and Mid-Year Examinations 2005 (M2005) were examined. The changes in the mean difference between group E and group C for the three school-based examinations were noted.

As reported earlier, group E and group C were non-equivalent groups[12] as there was significant difference between the two groups in the PSLE mathematics results and the overall PSLE scores, using the Mann Whitney U Test. Group E appeared to have higher mathematical abilities than group C, based on the PSLE Mathematics grades and hence it would be difficult to infer the influence of the interventions implemented in the classroom, in addition to the presence of other variables. To begin, the data collected was first tested for normality using the Kolmogorov–Smirnov test, as indicated in Table 8.3.

Table 8.3. The Kolmogorov-Smirnov Test for Group E and Group C

	Group (No of students)	Mean	SD	z	p	Accept H_0
M(2004)	E (40)	73.86	11.782	0.569	0.903	Normal
	C(40)	64.00	11.557	0.491	0.969	Normal
F(2004)	E(40)	83.01	9.621	0.684	0.738	Normal
	C(40)	74.64	14.771	1.169	0.130	Normal
M(2005)	E (40)	68.81	14.592	0.623	0.833	Normal
	C(41)	56.83	14.742	0.567	0.905	Normal

As the distributions of the sets of data are normal, an independent sample t-test was used to compare the sets of data: M(2004), F(2004) and M(2005) between group E and group C. The results revealed that there was significant difference between group E and group C for the three semestral examinations. Table 8.4 shows the results of the test.

Table 8.4. Results of T-test between Group E and Group C

Exam	Mean Difference	t	df	Sig. (2-tailed)
M(2004)	9.862	3.780	78	**0.000***
F(2004)	8.375	3.005	78	**0.004***
M(2005)	11.977	3.669	74	**0.000***

It was noted that the mean difference between the two groups for M(2004) was approximately 9.7, but was reduced to 8.4 for F(2004) and widened to 12.0 in M(2005). The mean scores for the three semestral examinations for both group E and group C are indicated in Figure 8.1.

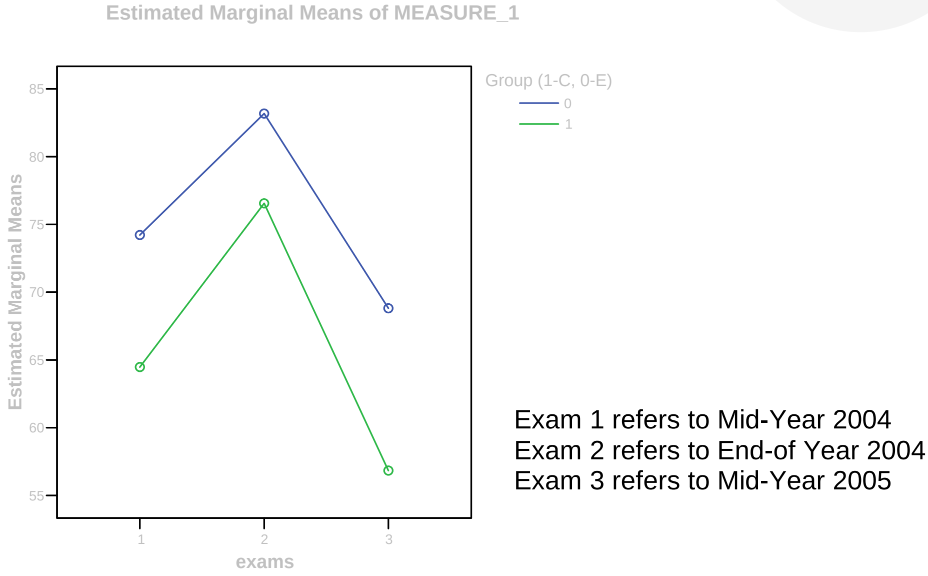


Figure 8.1. Mean value of school semestral examination results for group E and group C

One of the possible reasons for the narrowing of the gap when comparing M(2004) with F(2004), may be due to the initial teething problems that the mathematics teacher faced in integrating this new form of assessment strategy into her instructional classroom practice. Though Teacher E had been very positive about her involvement in this study, she had also on occasions in the early parts of the intervention expressed her reservations about using performance tasks because of the perceived ‘reduced’ curriculum time due to the ‘add-ons’ of performance tasks. Her presumptions had led her to conduct additional lessons for group E. In fact, it was also reported in the interview, conducted after the interventions that she had carried out a number of extra or remedial lessons with group E when they were in Secondary One. However, the number of ‘extras’ was reduced to only one or two lessons when group E was in Secondary Two. The following was extracted from the interview:

> **Teacher E:** I do not conduct extra lessons now, not on a regular basis. I mean only when it is required. I did quite a number when students were in Sec 1. But in Sec 2 maybe I conducted extra lessons only once or twice. And when it is nearer to exams, it was only for selected group of people, who really cannot follow. It’s combined with another class.

In another interview carried out with some average to low performing students, they revealed that Teacher E’s explanations did not match their ability level. In fact some students expressed that her lessons were pitched beyond their comprehension level, so they would rather seek help from their friends than approach her for assistance.

> **LP1:** If we don’t understand when we ask her, she would explain to us how to do it but we still don’t understand….. We feel that it’s very complicated. When teacher E explains, she does it in a complicated way. I am blurred. I don’t know what is she saying, but if my friend explains I find it easier.

> **LP2:** Her teaching is at a higher standard. When our friends teach us, it’s like we are in the same standard.

> **AA2:** Her teaching is too complicated for us to understand……. Most of the time, I would ask my friend. If I am on my own, I would read it over and over again but sometimes I still cannot understand.

In addition, one student, AA2 also perceived her to be strict. He said: "*Teacher E is quite strict*."

Teacher E was aware that they sought their friend's help in the learning of mathematics but attributed their actions to the strong camaraderie which existed among the students that helped to provide the support to overcome their learning difficulties. In the interview, she said: "*Even though there are times when they do not understand, they ask one another. They learn from their friends and they are certainly not the weaker ones. They are quite bright."*

To understand more about Teacher E and group E, the researcher conducted a follow-up interview with the school administrator who confirmed the findings that Teacher E in her overly strict ways of handling her students and her no nonsense style of classroom management was sometimes misconstrued by her students as aloof and unapproachable. In contrast, the warm and friendly disposition of the mathematics teacher in group C helped to put students at ease in her presence. Triangulating the information with the interviews, it appeared that the difference of 12 points between group E and group C, in the examinations scores of M(2005) which rose from 8.4 points in F(2004) may not be to a large degree, attributed to the influence of the teacher.

Interestingly, despite the presence of Teacher E who seemed somehow unapproachable, according to the interview with the school administrator and whose pedagogical approach did not match the ability of some students in group E as revealed in the interview with the students, the difference between the mean scores of group E and group C increased in the semestral examination from 8.4 points, F(2004) to 12 points, M(2005). Over-riding the effect of the teacher, the widening of the gap between the mean scores of group E and C, in favour of group E, seemed to suggest the positive influence of the interventions on the academic performance of the students in group E. However, caution needs to be taken in interpreting the results as the design of the study includes a non-equivalent comparison group as well as the presence of other possible variables.

8.4.2 Results on students' affects

There is a common perception that success or failure in solving mathematics problems often depends very much on the requisite content knowledge. However knowing the procedures, algorithms and the appropriate formula is not sufficient to guarantee success in problem- solving. Other affective factors such as feelings of anxiety, frustration, confidence and the pleasure that one feels while working on the task and even the beliefs one holds, influence the outcome of one's performance (Lester, Garofalo & Kroll, 1989). These affective factors, though not explicitly addressed in a typical classroom instruction, are nonetheless important aspects of mathematics teaching and learning.

In this sub-study, to address the second research question on the use of authentic and/or open-ended performance tasks, resulting in potential changes in the affective aspects of the students, a pre-questionnaire survey was administered to both group E and group C, prior to the intervention programme, to establish an initial point of reference on students' attitude towards mathematics. A post-questionnaire survey was conducted after the intervention period to gauge the changes in the affects of the students.

Table 8.5 shows the response rate of students in group E and group C in the pre- and post-questionnaire survey.

Table 8.5. Response rate of the pre- and post-questionnaire surveys

Group/Survey		No. of students who responded	Total no of students	Percentage
E	Pre-survey	40	40	100%
	Post-survey	39	40	97.5%
C	Pre-survey	40	40	100%
	Post-survey	34	40	85%

Analysis for part A of the pre- and post-questionnaire survey is based on the four domains:

- general perception towards mathematics and the learning of mathematics,
- anxiety in learning mathematics,
- perception of one's mathematical competence and
- belief about the usefulness of mathematics.

In order to examine the changes in the students' response to the question items, between group E and group C, prior to and after the intervention period, numerals ranging from 1 to 9 were assigned to the nine-point Likert-type scale - with '1' as 'disagree totally' and '9' as 'agree totally' in the questionnaire surveys. The number of responses on the scale for each question item in the questionnaire survey is converted to percentages to better reflect the changes. In addition, the mean value of the response for each item is calculated using the assigned numbers. Two statistical tests were used in the analysis:

- A Mann-Whitney U test was applied to verify the significant difference between group E and group C for the two questionnaire surveys;
- The Wilcoxon Signed Ranks test was used to confirm the significance in the difference between the pre- and post-questionnaire survey within each group. The '*z*' and '*p*' value for the test was indicated below the item statement, in all the tables detailing the percentage of students' response to question items.

In addition, the analysis of the question items for Part A of both surveys was triangulated with data from the interviews.

General perception

Six question items in the questionnaire survey: Q1, Q5, Q9, Q13, Q16 and Q19 were designed to examine the students' general views towards the learning of mathematics.

Table 8.6 below shows the response of students in group E to the six question items for this domain. The shaded rows[13] in Table 8.6 refer to the negatively worded statements. For example Item Q5 '*Mathematics is hard for me*'. A lower mean value would indicate a more favourable response to such items.

1. Group E

By comparing the percentages of students' response to the pre- and post-questionnaire survey, the data revealed that students in group E in general had become less positive in their responses to all the six items.

In Item Q5, by combining all the percentages starting from '*Agree a little*' to '*Agree totally*', 17.5% of the students in group E in the pre-questionnaire survey, agreed with the statement that '*mathematics is hard for me*'. However in the post-questionnaire survey, the percentage who agreed with the statement in Item Q5, increased to 33.5%. This implied that students felt mathematics was more difficult as compared to prior to the intervention of the new assessment strategy. The same manner of totalling the percentages from '*Disagree totally' to 'Disagree a little*' is adopted to represent the broad category of students who generally disagree with the statement in the question item. The data in the pre-questionnaire survey revealed that most students in group E enjoyed solving mathematics problems (77.5%) and felt that mathematics was interesting (82.1%). However in the post-questionnaire survey only 61.5% of students in group E continued to enjoy doing mathematics with 59% of students in group E agreed with the statement that 'mathematics is interesting to me'.

Table 8.6. Distribution of students' response to items on general perception for Group E

General perception towards Mathematics and learning of Mathematics												
Group E			DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q1	I enjoy doing mathematics. (z = -2.171, p =0.030*)	Pre(40)	5.0%	0.0%	2.5%	5.0%	10.0%	0.0%	20.0%	20.0%	37.5%	7.20
		Post(39)	12.8%	7.7%	5.1%	2.6%	10.3%	7.7%	17.9%	20.5%	15.4%	5.79
Q5	Mathematics is hard for me. (z = -2.073, p = 0.038*)	Pre(40)	22.5%	27.5%	12.5%	15.0%	5.0%	2.5%	5.0%	7.5%	2.5%	3.33
		Post(39)	2.6%	15.4%	30.8%	5.1%	12.8%	10.3%	10.3%	10.3%	2.6%	4.49
Q9	Mathematics is interesting to me. (z = -3.176, p = 0.001*)	Pre(39)	5.1%	0.0%	0.0%	10.3%	2.6%	7.7%	20.5%	10.3%	43.6%	7.23
		Post(39)	7.7%	5.1%	15.4%	2.6%	10.3%	12.8%	28.2%	15.4%	2.6%	5.46
Q13	I don't have good feelings about math (z = -1.699 , p = 0.089)	Pre(40)	42.5%	7.5%	7.5%	5.0%	22.5%	10.0%	2.5%	0.0%	2.5%	3.13
		Post(39)	12.8%	10.3%	20.5%	10.3%	17.9%	5.1%	15.4%	2.6%	5.1%	4.31
Q16	I like spending time on studying mathematics. (z = -2.783, p = 0.005*)	Pre(40)	2.5%	5.0%	5.0%	2.5%	10.0%	17.5%	17.5%	17.5%	22.5%	6.57
		Post(39)	5.1%	7.7%	7.7%	12.8%	28.2%	12.8%	12.8%	7.7%	5.1%	5.10
Q19	I don't like to attend math lessons. (z = -3.433, p = 0.001*)	Pre(40)	42.5%	12.5%	12.5%	7.5%	10.0%	5.0%	2.5%	5.0%	2.5%	2.95
		Post(39)	7.7%	10.3%	7.7%	2.6%	28.2%	15.4%	10.3%	7.7%	10.3%	5.21

Note. *less than 0.05 level of significance for two tail-test. DT: Disagree totally; DAL: Disagree a lot; D: Disagree; DLL: Disagree a little; N: Neither disagree nor agree; ALL: Agree a little; A: Agree; AAL: Agree a lot; AT: Agree totally. (The same notations apply to the tables below)

Three quarters of the students in group E responded that they '*like spending time on studying mathematics*' in the pre-questionnaire survey but less that two-fifths retained the same view in the post-questionnaire survey. For Item Q19, only 15% responded in agreement with the statement that they '*don't like to attend math lessons*' but the percentages rose to 43.7% in the post-questionnaire survey.

The Wilcoxon Signed Ranks test is used to detect the significant difference between the pre- and post-questionnaire surveys of group E, for the 6 question items under the domain of general perception towards mathematics and the learning of mathematics. With the exception of Item Q13, the test revealed that there was significant difference between the pre- and post-questionnaire for all the other items on general perception towards Mathematics and the learning of Mathematics.

In other words, group E has become less positive in their general perception towards mathematics and the learning of mathematics as compared to eighteen months ago before the interventions had taken place. The '*z'* and *'p'* value is indicated in Table 8.6 below the item statement.

2. Group C

Similarly, about three quarters of the students in group C responded that '*mathematics is very interesting*' (75%) and they '*enjoy solving mathematics problems*' (72.5%) in the pre-questionnaire survey. However in the post-questionnaire survey, the percentage of students in group C who found the subject interesting declined to 52.9% and only 50% of the students took pleasure in doing mathematics. The percentage of students in group C who agreed that '*mathematics is hard*' rose from 30% in the pre-questionnaire survey to 61.7% in the post-questionnaire survey. The percentage increase for group C (31.7%) was almost twice that of group E (16%) for this item.

Table 8.7 below shows the response of group C to the six question items for this domain.

Table 8.7. Distribution of students' response to items on general perception for Group C

General Perception towards Mathematics and the learning of Mathematics												
	Group C		DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q1	I enjoy doing mathematics (z = -2.838 p =0.005*)	Pre(40)	2.5%	2.5%	7.5%	7.5%	7.5%	5.0%	27.5%	20.0%	20.0%	6.60
		Post(34)	8.8%	11.8%	8.8%	14.7%	5.9%	14.7%	26.5%	8.8%	0.0%	4.91
Q5	Mathematics is hard for me. (z = -3.152, p = 0.002*)	Pre(40)	15.0%	12.5%	20.0%	10.0%	12.5%	17.5%	5.0%	5.0%	2.5%	4.05
		Post(34)	2.9%	8.8%	8.8%	5.9%	11.8%	23.5%	17.6%	14.7%	5.9%	5.65
Q9	Mathematics is interesting to me. (z = -2.500, p = 0.012*)	Pre(40)	2.5%	0.0%	10.0%	7.5%	5.0%	7.5%	25.0%	7.5%	35.0%	6.83
		Post(34)	2.9%	2.9%	17.6%	2.9%	20.6%	14.7%	26.5%	8.8%	2.9%	5.47
Q13	I don't have good feelings about mathematics (z = -1.30, p = 0.193)	Pre(40)	20.0%	17.5%	22.5%	5.0%	12.5%	7.5%	7.5%	2.5%	5.0%	3.68
		Post(34)	2.9%	11.8%	17.6%	14.7%	23.5%	23.5%	2.9%	2.9%	0.0%	4.41
Q16	I like spending time on studying mathematics. (z = -3.025, p = 0.002*)	Pre(40)	2.5%	2.5%	7.5%	12.5%	15.0%	17.5%	17.5%	12.5%	12.5%	5.95
		Post(34)	8.8%	5.9%	20.6%	11.8%	20.6%	14.7%	17.6%	0.0%	0.0%	4.44
Q19	I don't like to attend math lessons. (z = -3.747, p = 0.000*)	Pre(40)	20.0%	17.5%	20.0%	10.0%	20.0%	7.5%	5.0%	0.0%	0.0%	3.35
		Post(34)	2.9%	5.9%	8.8%	11.8%	26.5%	17.6%	8.8%	8.8%	8.8%	5.38

Note. *less than 0.05 level of significance for two-tailed test.

About 60% of students in group C responded that they '*like spending time on maths*' in the pre-survey but only 32.3% retained the same view in the post-survey. For Item Q19, 12.5% of students in group C agreed that they '*don't like to attend maths lessons*' but the percentage increased to 44% in the post- questionnaire survey.

Similarly, the Wilcoxon Signed Ranks test revealed that significant difference existed, between the pre- and post-questionnaire of group C for all the items with the exception of Item Q13.

It seemed that group C, just like group E, had become less positive towards mathematics and mathematics learning, as compared to 18 months ago. The '*z*' and '*p*' values were shown in Table 8.7 below the item statement.

3. Comparison of group E with group C

Both group E and group C had very positive general perceptions towards the learning of Mathematics in the pre-questionnaire survey. A Mann Whitney U test applied between the two groups did not register any significant difference for all the items in this domain in the pre- questionnaire survey. This implied that prior to the interventions, group E and group C had very similar perceptions towards Mathematics and the learning of Mathematics. In the post-questionnaire survey, significant difference was only detected for Item Q5, '*Mathematics is hard for me*', when a Mann Whitney U test was performed on the two groups. Table 8.8 shows the results of the Mann Whitney U test between group E and group C for the two questionnaire surveys.

Table 8.8. Comparison of general perception between Group E and Group C

Item		Pre-survey					Post-survey				
		Mean E(40)	Mean C(40)	U	z	Sig (2 –tail)	Mean E(39)	Mean C(34)	U	z	Sig (2 –tail)
Q1	I enjoy doing mathematics.	7.20	6.60	632.5	-1.650	0.099	5.79	4.91	502.0	-1.798	0.072
Q5	Mathematics is hard for me.	3.33	4.05	629.5	-1.660	0.097	**4.49**	**5.65**	**467.5**	**-2.185**	**0.029***
Q9	Math is interesting to me.	7.23	6.83	686.5	-0.953	0.341	5.46	5.47	637.5	-0.287	0.774
Q13	I don't have good feelings about mathematics.	3.13	3.68	672.0	-1.258	0.208	4.31	4.41	625.0	-0.425	0.671
Q16	I like spending time on studying mathematics.	6.57	5.95	643.0	-1.528	0.127	5.10	4.44	548.5	-1.283	0.199
Q19	I don't like to attend math lessons.	2.95	3.35	651.5	-1.460	0.144	5.21	5.38	647.0	-0.180	0.858

Note. *less than 0.05 level of significance for two-tailed test.

Webb (1992, p. 580) noted that as students grow older, there is a tendency to develop the notion, 'mathematics must be difficult and unfamiliar'. This trait could be observed in both group E and group C, but the increase from 30% in the pre-survey to 61.7% in the post-survey for group C was more acute than the increase for group E, from17.5% in the pre-survey to 33.5% in the post-survey. The large gap of about 30% between group E (33.5%) and group C (61.7%) in the post-survey for the question Item Q5, '*Mathematics is hard for me*' appeared to the researcher, to have been attributed to the exposure of group E to the performance tasks carried out in the study. The interventions implemented out over a period of eighteen months might have helped the students in group E to acquire problem solving skills in handling non-routine and open-ended problems, thereby fostering the positive perception that they were able to cope with difficult mathematics problems.

Triangulating this information with the data obtained through the interviews, the two low performing students revealed that they found the later few performance tasks much easier. LP2 said: "Oh *yes, for the later part of the programme, it's easier*". When the researcher sought clarification from the students if they were referring to the later part of the performance tasks or the entire intervention programme, LP1 reaffirmed: "*Yes, the later few tasks are much easier*". The responses from these students in the interview seemed to reaffirm the above point that the ten interventions implemented over a period of eighteen months had helped to nurture the positive views that they were able to cope with challenging problems.

Anxiety in the learning of maths

In this sub-study, mathematics anxiety refers to feelings of tension and nervousness that interferes with the mathematical performance within the context of school mathematics. Some students have negative attitudes towards mathematics that are often reflected in feelings of uneasiness when they are required to operate mathematically.

As mentioned earlier, a total of six question items: Q2, Q6, Q10, Q17 and Q20 in the survey were designed to examine students' level of anxiety when solving mathematics problems.

1. Group E

In the pre-questionnaire survey, 67.5% held the opinion that they were '*never under a terrible strain in a maths class*' but in the post-questionnaire the response for this item declined to 41%.

Similarly the percentage for those who reported that they were not fearful of doing mathematics fell slightly by 8.4%, from 75% in the pre-survey to 66.6% in the post-survey. 25% of students in group E felt lost in solving mathematics problems in the pre-survey but the percentages increased marginally to 30.9%. Interestingly, the students who felt 'nervous to even think about having to do a mathematics problem' reduced by half from 25% in the pre-survey to 12.5% in the post-survey. However those who felt that they lacked the clarity of thoughts when working on mathematics almost doubled form 17.5% in the pre-survey to 28.3% in the post-survey.

Table 8.9 below shows the responses of group E to the six question items for this domain on anxiety for both the pre- and post-questionnaire survey.

Table 8.9. Distribution of students' response to items on feelings of anxiety for Group E

	Feelings of Anxiety											
	Group E		DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q2	I am never under a terrible strain in a math class. (z = -2.841 p = 0.004*)	Pre(40)	7.5%	2.5%	2.5%	5.0%	15.0%	12.5%	17.5%	27.5%	10.0%	6.23
		Post(39)	7.7%	12.8%	15.4%	12.8%	10.3%	17.9%	10.3%	12.8%	0.0%	4.64
Q6	I am not afraid of doing maths. (z =-1.439, p = 0.150)	Pre(40)	5.0%	2.5%	5.0%	2.5%	10.0%	5.0%	10.0%	22.5%	37.5%	7.03
		Post(39)	2.6%	0.0%	7.7%	5.1%	17.9%	5.1%	35.9%	12.8%	12.8%	6.36
Q10	I am unable to think clearly when doing Maths (z =-1.710, p = 0.087)	Pre(40)	27.5%	22.5%	12.5%	10.0%	10.0%	10.0%	2.5%	2.5%	2.5%	3.20
		Post(39)	7.7%	10.3%	33.3%	10.3%	10.3%	12.8%	10.3%	2.6%	2.6%	4.13
Q14	I feel lost when trying to solve math problems. (z = -1.758, p = 0.079)	Pre(40)	22.5%	20.0%	17.5%	10.0%	5.0%	15.0%	7.5%	0.0%	2.5%	3.45
		Post(39)	12.8%	7.7%	25.6%	7.7%	17.9%	10.3%	2.6%	5.1%	10.3%	4.38
Q17	It makes me nervous to even think about having to do a maths problem. (z = -0.540, p = 0.589)	Pre (40)	35.0%	17.5%	7.5%	5.0%	10.0%	15.0%	7.5%	0.0%	2.5%	3.28
		Post(39)	12.8%	12.8%	33.3%	15.4%	12.8%	2.6%	10.3%	0.0%	0.0%	3.51
Q20	I have a lot of confidence when it comes to maths (z = -2.227, p = 0.026*)	Pre (40)	2.5%	5.0%	2.5%	10.0%	2.5%	15.0%	15.0%	17.5%	30.0%	6.78
		Post(39)	2.6%	5.1%	17.9%	5.1%	15.4%	20.5%	15.4%	7.7%	10.3%	5.49

Note. *less than 0.05 level of significance for two-tailed test.

From Table 8.9, we can see that 77.5% of students in group E claimed that they '*have a lot of confidence when it comes to mathematics*' in the pre-survey but this declined by approximately 24% to 53.9% for the post-survey. A within group comparison using the Wilcoxon Signed Rank test for the pre- and post-questionnaire survey revealed that there is significant difference for Item 2 and Item 20 for group E.

2. Group C

For Item Q2, 62.5% in group C reported that '*they were never under a terrible strain in a maths class*' but the percentage reduced drastically to 23.5% in the post-survey. The percentage of students who held the opinion that they were not fearful of mathematics reduced from 70% to 58.8%.

Table 8.10 below shows the responses of group C to the six question items for this domain on anxiety for both the pre- and post-questionnaire survey.

Table 8.10. Distribution of students' response to items on feelings of anxiety for Group C

	Feelings of anxiety											
Group C			DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q2	I am never under a terrible strain in a math class. (z = -3.403, p = 0.001*)	Pre(40)	5.0%	2.5%	10.0%	5.0%	15.0%	17.5%	22.5%	15.0%	7.5%	5.85
		Post(34)	8.8%	11.8%	35.3%	11.8%	8.8%	5.9%	14.7%	2.9%	0.0%	3.91
Q6	I am not afraid of doing maths. (z =-1.250, p = 0.211)	Pre(40)	7.5%	2.5%	0.0%	10.0%	10.0%	7.5%	15.0%	20.0%	27.5%	6.60
		Post(34)	2.9%	2.9%	5.9%	17.6%	11.8%	8.8%	35.3%	2.9%	11.8%	5.85
Q10	I am unable to think clearly when doing Maths. (z =-0.583, p = 0.560)	Pre(40)	10.0%	12.5%	15.0%	12.5%	15.0%	10.0%	10.0%	10.0%	5.0%	4.60
		Post(34)	5.9%	5.9%	14.7%	5.9%	32.4%	17.6%	11.8%	2.9%	2.9%	4.85
Q14	I feel lost when trying to solve math problems. (z =-1.767, p = 0.077)	Pre(40)	12.8%	10.3%	10.3%	15.4%	15.4%	17.9%	10.3%	5.1%	2.6%	4.46
		Post(34)	2.9%	8.8%	8.8%	14.7%	11.8%	29.4%	14.7%	5.9%	2.9%	5.18
Q17	It makes me nervous to even think about having to do a maths problem. (z = -0.980, p = 0.327)	Pre(40)	22.5%	2.5%	15.0%	10.0%	22.5%	10.0%	7.5%	7.5%	2.5%	4.20
		Post(34)	2.9%	2.9%	29.4%	5.9%	23.5%	20.6%	5.9%	5.9%	2.9%	4.76
Q20	I have a lot of confidence when it comes to maths. (z =-2.659, p = 0.008*)	Pre(40)	0.0%	2.5%	12.5%	12.5%	30.0%	5.0%	7.5%	17.5%	12.5%	5.78
		Post(34)	8.8%	2.9%	23.5%	20.6%	20.6%	5.9%	11.8%	5.9%	0.0%	4.35

Note. *less than 0.05 level of significance for two-tailed test.

Similarly, the percentage of students who agreed that they felt lost in solving mathematics problems rose from 35.9% to 52.9% while the percentage of students in group C who held the same opinion that they were '*unable to think clearly when doing mathematics*', maintained at 35% for both the pre- and post-survey. In particular for Item 20, the percentage of students who felt they '*have a lot of confidence in mathematics*' fell from 42.5% to 23.6%. Using the Wilcoxon Signed Rank test for a within group comparison for group C on the pre- and post- questionnaire survey revealed that there is significant difference for Item 2 and Item 20. The results were similar to that of group E.

3. Comparison of group E with group C

Although the feelings of anxiety rose for both groups comparing the pre- and post-questionnaire survey, group E as compared to group C continued to exhibit a stronger sense of self-confidence and appeared to feel less anxious in learning the subject in the post-questionnaire survey. In particular for Item Q17, the percentage who held the view about feeling nervous to even think about having to solve a mathematics problem reduced from 25% in pre-survey to almost half in the post-survey for group E.

On the contrary, the response of group C for Item Q17, increased from 27.5% to 35.3%. The low level of anxiety in the learning of mathematics was also evident in group E's response to Item 20, with 77.5% in the pre-survey who concurred with the statement, 'I have a lot of confidence when it comes to maths' and 53.9% in the post-survey. In contrast, the percentage who agreed with the statement for group C was 42.5% for the pre-questionnaire survey but was reduced to 23.6% in the post- survey.

Table 8.11 below shows the mean value for both groups E and C in their response to the questionnaire on the feelings of anxiety and the results of the Mann Whitney U test between the two groups.

Table 8.11. Comparison of feelings of anxiety between Group E and Group C

Item		Pre-survey					Post-survey				
		Mean E(40)	Mean C(40)	U	z	Sig (2 –tail)	Mean E(39)	Mean C(34)	U	z	Sig (2 –tail)
Q2	I am never under a terrible strain in a math class.	6.23	5.85	695.5	-1.019	0.308	4.64	3.91	537.0	-1.411	0.158
Q6	I am not afraid of doing maths.	7.03	6.60	698.0	-1.006	0.315	6.36	5.85	558.0	-1.193	0.233
Q10	I am unable to think clearly when doing maths	**3.20**	**4.60**	**521.5**	**-2.706**	**0.007***	4.13	4.85	516.5	-1.645	0.100
Q14	I feel lost when trying to solve math problems.	**3.45**	**4.46**	**575.0**	**-2.030**	**0.042***	4.51	5.18	502.0	-1.501	0.133
Q17	It makes me nervous to even think about having to do a maths problem	3.28	4.20	631.5	-1.649	0.099	**3.51**	**4.76**	**414.5**	**-2.806**	**0.005***
Q20	I have a lot of confidence when it comes to maths	**6.78**	**5.78**	**571.0**	**-2.230**	**0.026***	**5.49**	**4.35**	**464.5**	**-2.220**	**0.026***

Note. *less than 0.05 level of significance for two-tailed test.

Hembree (1990) in a meta-analysis based on 151 studies confirmed that mathematics anxiety is related to poor performance in mathematics or low mathematical achievement. In other words, a reduction in mathematics anxiety which relates inversely to positive attitudes such as self-confidence, induce higher mathematics achievement. However for group E, the feelings of anxiety might be a positive indication that students were more concerned about learning the subject as they progressed academically. It was also noted by Teacher E that the students in group E were motivated to do well academically. She said, "*they are quite concern with their work.*"

Prior to the interventions, there was significant difference between the two groups for Item Q10, Item Q14 and Item Q20. In the post-questionnaire survey, significant difference was only registered for two questions – Items Q17 and Item Q20.

On the whole, group E appeared to feel less anxious as compared to group C in both the pre- and post-survey. As mathematics anxiety is inversely related to self-confidence, the feelings of being more confident towards the learning of mathematics for group E, as compared to group C in the post-survey seemed to the researcher, to be ascribed to the use of authentic and/or open-ended performance tasks for the past eighteen months. As these tasks were more challenging and demanding in comparison to routine textbook questions, students developed self-confidence when they were provided with opportunities to engage in higher-order thinking. The classroom presentation and discussion might have also helped to reinforce the positive affects when students communicate their solutions to the entire class.

Perception of mathematical competence
Confidence in learning mathematics is an important affective factor and plays a crucial role in students' mathematical achievement as well as the ability to solve non- routine problems (Hart & Walker, 1993). In this study, six question items were designed to examine the students' perception of their mathematical competence: Q3, Q7, Q11, Q15, Q21 and Q22.

1. Group E

In the pre-survey, 90% of the students in group E felt assured of achieving good grades in mathematics and 85% felt that they could learn well in the Mathematics subject. However, in the post-questionnaire survey, the percentage fell to 79.5% and 69.5 % respectively for these two items.

Table 8.12 below shows the response of group E for this domain on perception of mathematical competence for both the pre- and post- questionnaire survey.

Interestingly, 22.5% of students in group E agreed that they were '*not good in mathematics*' in the pre-survey and in the post-survey, the percentage increased approximately by 10% to 33.3%. Similarly, 12.5% of students in group E reported that they '*do not think that they can do well in mathematics*' and the percentage increased marginally by 8% to 20.5% in the post-survey.

Table 8.12. Distribution of students' response to items on Math competence for Group E

	Perception of mathematical competence											
	Group E		DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q3	I am sure I can learn maths well. (z = -1.236, p = 0.216)	Pre(40)	2.5%	2.5%	2.5%	5.0%	2.5%	5.0%	25.0%	20.0%	35.0%	7.28
		Post(39)	0.0%	2.6%	7.7%	0.0%	10.3%	20.5%	23.1%	20.5%	15.4%	6.67
Q7	I can get good grades in maths. (z = 0.150, p = 0.094)	Pre (40)	2.5%	0.0%	0.0%	2.5%	5.0%	17.5%	27.5%	20.0%	25.0%	7.20
		Post(39)	5.1%	2.6%	5.1%	7.7%	10.3%	2.6%	33.3%	20.5%	12.8%	6.36
Q11	I am not good at maths. (z = -1.728, p = 0.084)	Pre (40)	22.5%	25.0%	17.5%	2.5%	10.0%	7.5%	5.0%	5.0%	5.0%	3.50
		Post(39)	12.8%	5.1%	25.6%	7.7%	15.4%	5.1%	12.8%	5.1%	10.3%	4.62
Q15	I don't think I can do well in maths. (z= -2.675, p = 0.007*)	Pre (40)	45.0%	15.0%	12.5%	12.5%	2.5%	7.5%	0.0%	2.5%	2.5%	2.63
		Post(39)	12.8%	12.8%	25.6%	12.8%	15.4%	7.7%	5.1%	2.6%	5.1%	3.92
Q21	I like solving challenging maths problems. (z = -2.150, p = 0.032*)	Pre (40)	2.5%	10.0%	2.5%	5.0%	5.0%	10.0%	12.5%	15.0%	37.5%	6.80
		Post(39)	12.8%	10.3%	5.1%	5.1%	7.7%	20.5%	15.4%	10.3%	12.8%	5.36
Q22	I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself. (z = -2.425, p = 0.015*)	Pre(40)	30.0%	15.0%	10.0%	7.5%	25.0%	2.5%	5.0%	5.0%	0.0%	3.35
		Post(39)	5.1%	7.7%	23.1%	10.3%	30.8%	5.1%	2.6%	5.1%	10.3%	4.67

Note. *less than 0.05 level of significance for two-tailed test.

Table 8.12 also shows that for Item Q21 and Item Q22, in the pre-survey, three quarters of students in group E liked '*solving challenging mathematics problems*' while one eighth preferred to receive the solution to a difficult problem rather than having to solve it themselves. However, in the post- survey, the fraction of students who continue to take pleasure in solving challenging questions declined to three-fifths and the fraction of students in group E who preferred to receive the solution to a difficult problem rather than having to solve it themselves double to almost a quarter.

2. Group C

In comparison, in the pre-questionnaire survey, group C seemed to feel less confident, with 65% of the students assured of getting good grades and 77.5% deemed that they could learn the subject well.

Table 8.13 shows the response of group C students for this domain on perception of mathematical competence for both the pre- and post-questionnaire survey. As the table shows, in the post-questionnaire survey, the percentage of students who felt assured of good grades declined to 41.2% while the number who continued to feel that they could learn mathematics well, fell by 9.8% to 67.7%.

Table 8.13. Distribution of students' response to items on Math competence for Group C

	Perception of mathematical competence											
	Group C		DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q3	I am sure I can learn maths well. (z = -2.373, p = 0.018*)	Pre(40)	0.0%	0.0%	2.5%	7.5%	12.5%	5.0%	32.5%	17.5%	22.5%	7.00
		Post(34)	0.0%	2.9%	5.9%	8.8%	14.7%	29.4%	20.6%	11.8%	5.9%	6.00
Q7	I can get good grades in maths. (z = -2.120, p =0.034*)	Pre(40)	7.5%	2.5%	0.0%	10.0%	15.0%	12.5%	20.0%	20.0%	12.5%	6.15
		Post(34)	5.9%	5.9%	14.7%	11.8%	20.6%	20.6%	11.8%	5.9%	2.9%	4.91
Q11	I am not good at maths. (z = -1.178, p = 0.239)	Pre(40)	2.5%	10.0%	30.0%	7.5%	25.0%	10.0%	5.0%	2.5%	7.5%	4.50
		Post(34)	0.0%	11.8%	17.6%	11.8%	17.6%	11.8%	14.7%	5.9%	8.8%	5.12
Q15	I don't think I can do well in maths. (z = -1.710, p = 0.087)	Pre(40)	22.5%	7.5%	22.5%	10.0%	17.5%	10.0%	5.0%	2.5%	2.5%	3.70
		Post(34)	2.9%	5.9%	20.6%	20.6%	20.6%	14.7%	8.8%	2.9%	2.9%	4.62
Q21	I like solving challenging maths problems. (z = -3.446, p = 0.001*)	Pre (40)	2.5%	0.0%	2.5%	12.5%	20.0%	10.0%	20.0%	12.5%	20.0%	6.40
		Post(34)	8.8%	0.0%	17.6%	20.6%	14.7%	17.6%	8.8%	8.8%	2.9%	4.82
Q22	I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself. (z = -1.359, p = 0.174)	Pre (40)	20.0%	5.0%	15.0%	20.0%	30.0%	0.0%	5.0%	0.0%	5.0%	3.85
		Post(34)	8.8%	2.9%	14.7%	14.7%	26.5%	11.8%	14.7%	0.0%	5.9%	4.76

Note. *less than 0.05 level of significance for two-tailed test.

For the two negatively worded items, Q11 and Q15, the percentage of students who reported that they were not good in mathematics increased from 25% in the pre-survey to 41.2% in the post-survey. Comparing the pre- and post-survey, the increase in the number of students who felt that they could not do well in mathematics for both groups was marginal.

In the pre-questionnaire survey, for Item Q21 and Item Q22, five-eighths of students in group C '*like solving challenging mathematics problems*' while one tenth preferred to receive the solution to a difficult problem rather than having to solve it themselves. However, in the post-survey, the fraction in group C who continued to take pleasure in solving challenging questions dipped to almost three-eighths and the fraction who preferred to receive the solution to a difficult problem rather than having to solve it themselves tripled to slightly more than three-tenths.

3. Comparison of group E with group C

On the whole, in the pre-survey, group E appeared to exhibit a greater level of confidence in the learning of Mathematics as compared to group C. A Mann Whitney U test showed that there was significant difference between group E and group C for three of the six items: Item Q7, Item Q11 and Item Q15 in the pre-survey.

In the post-questionnaire survey, group E continued to exhibit a greater level of confidence in the learning of Mathematics as compared to group C. However, only Item Q7 remained significant when the comparison was made between the two groups in the post-survey using the Mann Whitney U test.

For both the pre- and post-questionnaire surveys, group E appeared to exhibit a greater sense of self-assurance about learning and achieving in mathematics, in comparison to group C. In other words, group E held strong views about their mathematical competence and continued to do so prior to and after the intervention period. Close to 80% in group E responded in the post-survey that they could learn mathematics well and about 70% felt that they could get good grades in mathematics after the intervention programme. Only 20.5% expressed doubts about doing well in mathematics. It seems that the exposure to performance tasks may have a positive influence on their perception of their own mathematical competence. Table 8.14 below shows the '*U*', '*z*' and '*p*' value of the Mann Whitney U test between group E and group C on the perception of mathematical competence for the pre- and post-questionnaire survey.

Table 8.14. Comparison of perception of Maths competence between Group E and Group C

	Item	Pre-survey					Post-survey				
		E(40)	C(40)	U	z	Sig (2 –tail)	E(39)	C(34)	U	z	Sig (2 –tail)
Q3	I am sure I can learn mathematics well.	7.28	7.00	676.0	-1.227	0.220	6.67	6.00	499.0	-1.845	0.065
Q7	I can get good grades in mathematics.	**7.20**	**6.15**	**579.0**	**-2.162**	**0.031***	**6.36**	**4.91**	**385.0**	**-3.109**	**0.002***
Q11	I am not good at mathematics.	**3.50**	**4.50**	**546.0**	**-2.478**	**0.013***	4.62	5.12	576.0	-0.972	0.331
Q15	I don't think I can do well in maths.	**2.63**	**3.70**	**550.0**	**-2.466**	**0.014***	3.92	4.62	508.0	-1.737	0.082
Q21	I like solving challenging math problems.	6.80	6.40	663.0	-1.341	0.180	5.36	4.82	557.5	-1.176	0.240
Q22	I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself.	3.35	3.85	700.0	-0.983	0.326	4.67	4.76	615.5	-0.535	0.593

Note. *less than 0.05 level of significance for two-tailed test.

In the pre-questionnaire survey, there was significant difference for three question items between group E and group C when the Mann Whitney U test was used: Item Q7, Item Q11 and Item Q15. However in the post- questionnaire survey, only Item 7 remained significantly different between the groups using the same test.

The reduction in the number of items being significantly different, prior to and after the intervention programme may seem that the level of confidence for group E has declined due to the implementation of performance tasks. However a closer examination of the two separate interviews with the students provided a different perspective of the influence of interventions on their perceived mathematical ability.

In an interview with the average-ability student AA1, he revealed that performance tasks though challenging, nevertheless were manageable with occasional help. However when the researcher enquired about extending the use of performance tasks to the next semester, student AA1 was apprehensive because he felt that there were too many things to learn in the forthcoming semester. The following was extracted from the interview:

> **AA1:** ...but when it comes to performance tasks, we use the basic knowledge to try to solve more challenging problems. I think it's tougher doing the performance tasks.
>
> **Researcher:** You found them more challenging but do you enjoy doing it?
>
> **AA1:** Yes, I like challenges... Actually I find the performance tasks fun and easy to learn. Nothing is really difficult.
>
> **Researcher:** You find them manageable?
>
> **AA1:** Quite manageable. Sometimes there are difficulties but I think we can manage with a little help.
>
> **Researcher:** Can I carry on with more performance tasks in the next semester?
>
> **AA1:** I have gotten used to doing performance tasks, after being exposed to so many and I think I am quite comfortable now... Of course I would like to have it. But it depends whether I can manage because now I have too many things to learn. But I think I can.

It appeared that the average-ability student somehow came to realise that he needed to improve his content knowledge and developed his skills in several areas after solving performance tasks for the past eighteen months. His sense of self-confidence in his own mathematical ability declined.

A separate interview with the two high-performing students, HP1 and HP2, revealed a different form of concern.

> **HP2:** Some questions have many answers and we are unable to find so many answers…..For example this task (Decorating the Kitchen), many of us do not even understand the question.
>
> **Researcher:** Did your teacher try to help you out with that?
>
> **HP1 & HP2:** Yes, she did.
>
> **HP2:** But we are still unable to obtain all the answers.

Both high performing students felt apprehensive because they were not confident of obtaining the complete solution with all the answers, though they were able to solve the performance tasks. It appeared to the researcher that both students who had been achieving in the traditional semestral examinations had high expectations of themselves. Their sense of confidence about their mathematical ability diminished when they realised that they would not be able to peak in their performance or they might lack certain mathematical conceptual understanding to provide a complete solution to the problem when performance tasks were in use.

Beliefs about usefulness of mathematics
Perceived usefulness of mathematics is an important affective factor. Students, who are confident in learning mathematics and value mathematics as an important subject with practical functions, that is necessary for their future development, would very likely persevere to achieve in mathematics. Four items were designed to measure the component on students' perception about the usefulness of mathematics.

1. Group E
In the pre-survey, almost all the students in group E (97.5%) held the belief that mathematics is important and useful. However in the post-questionnaire survey, approximately 80% of the students in group E held the same view. 90% of the students from group E in the pre-survey disagreed that '*studying mathematics is a waste of time*' but the percentage declined drastically to 61.6% in the post-survey. Similarly, 80% in the group E agreed that they would use mathematics a lot as an adult but the percentage reduced to 64.1% in the post-survey. Comparing both the pre- and post-questionnaire survey, there was significant difference in group E's response to all the four question items.

Table 8.15 presents the response of group E to items on beliefs about the usefulness of mathematics for both the pre- and post-questionnaire survey.

Table 8.15. Distribution of response to items on usefulness of Mathematics for Group E

Beliefs about usefulness of Mathematics												
Group E			DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q4	I believe maths is useful. (z = -2.371, p= 0.018*)	Pre(40)	0.0%	0.0%	0.0%	0.0%	2.5%	7.5%	15.0%	15.0%	60.0%	8.23
		Post(34)	0.0%	5.1%	2.6%	2.6%	7.7%	5.1%	25.6%	23.1%	28.2%	7.15
Q8	It is important to know maths nowadays. (z = -2.484, p = 0.013*)	Pre(40)	0.0%	0.0%	0.0%	0.0%	2.5%	12.5%	12.5%	20.0%	52.5%	8.08
		Post(34	0.0%	0.0%	5.1%	0.0%	15.4%	10.3%	20.5%	20.5%	28.2%	7.15
Q12	Studying mathematics is a waste of time. (z = -2.688, p =0.007*)	Pre(40)	60.0%	7.5%	12.5%	10.0%	5.0%	2.5%	0.0%	0.0%	2.5%	2.15
		Post(34)	23.1%	10.3%	17.9%	10.3%	23.1%	10.3%	5.1%	0.0%	0.0%	3.51
Q18	I will use maths a lot as an adult. (z = -2.720, p = 0.007*)	Pre(40)	0.0%	2.5%	5.0%	0.0%	12.5%	2.5%	12.5%	10.0%	55.0%	7.60
		Post(34)	2.6%	5.1%	2.6%	5.1%	20.5%	15.4%	15.4%	28.2%	5.1%	6.15

Note. *less than 0.05 level of significance for two-tailed test.

2. Group C
In the pre-questionnaire survey, almost all the students in group C felt that mathematics was important (95%) but only 87.5% agreed that it was useful. In the post-questionnaire survey, however, 85.3% felt that mathematics was important while 88.1% felt that mathematics was useful. Both percentages were slightly higher than that of group E. In the pre- questionnaire survey, 90% of the students in group C disagreed that studying mathematics was a waste of time but only 79.5% retained that view in the post-questionnaire survey. In the pre-survey, 82.5% held the opinion that they would use mathematics a lot as an adult but only 67.5% continued to hold that view in the post-survey. Table 8.16 shows the response of group C to the four items on beliefs about the usefulness of mathematics in both the pre- and post- surveys.

Table 8.16. Distribution of response to items on usefulness of Mathematics for Group C

	Group C		DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
			Beliefs about usefulness of Mathematics									
Q4	I believe maths is useful. (z = -2.183, p= 0.029*)	Pre(40)	0.0%	0.0%	0.0%	0.0%	5.0%	5.0%	15.0%	20.0%	55.0%	8.15
		Post(34)	0.0%	0.0%	2.9%	0.0%	11.8%	5.9%	35.3%	17.6%	26.5%	7.29
Q8	It is important to know maths nowadays. (z = -2.247, p = 0.025*)	Pre(40)	0.0%	0.0%	0.0%	0.0%	12.5%	5.0%	12.5%	25.0%	45.0%	7.85
		Post(34)	0.0%	2.9%	0.0%	0.0%	8.8%	8.8%	44.1%	17.6%	17.6%	7.12
Q12	Studying maths is a waste of time. (z = -1.861, p =0.063)	Pre(40)	45.0%	15.0%	25.0%	5.0%	7.5%	0.0%	0.0%	2.5%	0.0%	2.28
		Post(34)	26.5%	11.8%	29.4%	11.8%	11.8%	8.8%	0.0%	0.0%	0.0%	2.97
Q18	I will use maths a lot as an adult. (z = -2.147, p = 0.032*)	Pre(40)	0.0%	0.0%	0.0%	5.0%	12.5%	5.0%	25.0%	7.5%	45.0%	7.53
		Post(34)	2.9%	0.0%	2.9%	0.0%	26.5%	20.6%	23.5%	8.8%	14.7%	6.35

Note. *less than 0.05 level of significance for two-tailed test.

3. Comparison of group E with group C

There is no significant difference between the groups using the Mann Whitney U test for the pre- and post-questionnaire survey. However about 15% of students in group E felt that studying mathematics was a waste of time while only 8.8% in group C felt the same way. Table 4.15 shows the results of the Independent Sample Mann Whitney U test between group E and C, on the perception about the usefulness of mathematics.

Table 8.17. Comparison of beliefs about usefulness of Mathematics between Group E and Group C

	Item	Pre-survey					Post-survey				
		E(40)	C(40)	U	z	Sig (2 –tail)	E(39)	C(34)	U	z	Sig (2 –tail)
Q4	I believe mathematics is useful.	8.23	8.15	768.0	-0.034	0.731	7.15	7.29	651.0	-0.137	0.891
Q8	It is important to know maths nowadays.	8.08	7.85	733.0	-0.069	0.489	7.15	7.12	630.0	-0.375	0.708
Q12	Studying maths is a waste of time.	2.15	2.28	718.0	-0.857	0.391	3.51	2.97	556.5	-1.200	0.230
Q18	I will use mathematics a lot as an adult.	7.60	7.53	735.5	-0.667	0.505	6.15	6.35	644.5	-0.208	0.835

Note. *less than 0.05 level of significance for two-tailed test.

In the interview with two high performing students from group E, they articulated that mathematics involved the direct application of formula and was easy. HP1 said: "*The questions in the textbook are more straight-forward*" while HP2 expressed: "*The questions in the textbook can be solved easily if we know the formula*".

In another interview, the average ability student, AA1 felt that the mathematics taught in the classroom was merely basic mathematics. He said: "*I think the math problem that we do in class is the basics of mathematics*."

Menon (2000) noted that students were drilled to solve numerous similar routine questions by applying familiar algorithms and following prescribed procedures. It appeared to the researcher that some students held similar notion that learning mathematics and solving mathematics problems necessitated only the application of familiar formula found in the textbook. Hence students who believed that they had no difficulties in applying procedural knowledge in mathematics and were confident of achieving good grades might feel that '*studying maths is a waste of time*'. These students would very likely prefer to devote more time to other subjects.

To conclude this section, efforts to integrate the new assessment strategy is a major goal of this study and to achieve this goal, the researcher needs to investigate the affective aspects of the students, related to its use in the mathematics classroom. In general, the attitudes towards the learning of mathematics declined for the groups, E and C as they progressed from Secondary One to Secondary Two. There are several possible reasons to explain the dip in affects when comparing the pre- and post-survey for both groups. According to Hart and Walker (1993), the decline in self-confidence as experienced by students when they progressed from elementary to the middle school, is a natural process of growing-up. Similarly, Strutchens (1999) also noted that students' confidence as well as the belief that they can do well in mathematics decreased over the years. Hence, the overall decline in the 4 dimensions: general attitude, anxiety, perception of their own mathematical competence, perceptions about the usefulness of mathematics for both the groups, E and C, is regarded as a natural phenomenon in their learning journey.

Hart and Walker (1993) also noted that students' perception of their own competence developed during the middle grades which they carry to high school, has long term implications on their mathematics achievement. Hence a nurturing school environment, care and concern from teachers and opportunities for students to demonstrate and realise their potential is extremely important to the affects and achievement of the students.

Despite the literature review which implied that the student self-confidence declined as they progressed from elementary levels, the findings seemed to suggest that group E remained positive in their disposition towards learning mathematics, prior to and after the intervention period. This is also evident in the number of question items that were significant in the post-questionnaire survey between group E and group C, in favour of group E:

- Mathematics is hard for me,
- It makes me nervous to even think about having to do a maths problem,
- I have a lot of confidence when it comes to maths,
- I can get good grades in mathematics.

It appeared that one possible contributing factor to explain group E's level of confidence in mathematics after the intervention period, as compared to group C, was the exposure of group E to performance tasks in the mathematics classroom. The authentic and/or open-ended performance tasks provided opportunities for students to engage in higher-order thinking in solving non-routine questions that relates to the real-life situation. However the extent of the changes that can be attributed to the use of the new assessment strategy needs a more in-depth study.

8.4.3 Results on student perception on performance tasks

Ten performance tasks (authentic and/or open-ended tasks) were implemented within the timeframe of eighteen months in the study. To examine the students' views on the performance tasks after the intervention programme, a total of sixteen questions were designed in part B of the post-questionnaire survey. The discussion centred on the following three main aspects:

- the students' perception of on the use performance tasks

- the challenges they face with this new strategy
- the perceived benefits derived from working on performance tasks

Similar to Part A of the questionnaire survey, numerals ranging from 1 to 9 were also assigned to the nine-point Likert-type scale with '1' as 'disagree totally' and '9' as 'agree totally'. The mean value of the response for each item was calculated using the assigned numbers. The number of responses on the scale for each item in the questionnaire survey was also converted to percentages. In addition, the analysis of the question items for Part B of post-questionnaire survey was triangulated with data from the interviews.

Students' views on performance tasks
Five items in part B of the post-questionnaire survey were designed to examine the views of group E on the use of performance tasks.

By combining all the percentages starting from 'A*gree a little'* to 'A*gree totally'*, it was observed that about 26% of the students in group E liked the feature on multiple solutions, while 56% responded that they enjoyed working on questions with multiple approaches. Approximately 26% felt that they were good at solving performance tasks and only 12.8% liked more performance tasks in their lessons. Table 8.18 shows the response of group E to the individual items.

Table 8.18. Distribution of response to views on performance tasks (Group E)

	Item (No of students =39)	DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q23	I like to solve maths questions which have more than one correct answer.	17.9%	5.1%	23.1%	12.8%	15.4%	7.7%	10.3%	2.6%	5.1%	4.10
Q26	I like to do maths questions which could be solved using different methods.	7.7%	2.6%	7.7%	12.8%	12.8%	12.8%	28.2%	7.7%	7.7%	5.56
Q30	I like to do maths questions which involve the real world	7.7%	2.6%	0.0%	12.8%	25.6%	15.4%	17.9%	7.7%	10.3%	5.64
Q34	I am good at doing maths performance tasks.	17.9%	2.6%	10.3%	17.9%	25.6%	17.9%	5.1%	2.6%	0.0%	4.18
Q36	I would like to have more maths performance tasks for my maths lessons	33.3%	2.6%	12.8%	7.7%	30.8%	5.1%	7.7%	0.0%	0.0%	3.46

In the interviews conducted with six students from group E, all revealed that they found the performance tasks more challenging as compared to routine problems in the textbooks – which they deemed only required the application of procedural knowledge. The students also acknowledged that the open-nature of the problem demanded them to spend more time thinking about the different approaches and situations as they worked out multiple solutions.

HP2: The questions in the textbook can be solved easily if we know the formula.

HP1: The questions in the textbook are more straight-forward. These questions (performance tasks) are more challenging, and take more time to solve.

AA1: I think the math problem that we do in class is the basics of mathematics. But when it comes to performance tasks, we use the basic knowledge to try to solve more challenging problems. I think it's tougher doing the performance tasks.

AA2: Yes we can have more ways to solve problems.

LP2: These types of questions (performance tasks) are very open. You have a lot of answers, not only one answer but more.

Slightly more than 50% of the students liked mathematics problems that relate to real world situations. Triangulating this information with the interviews, LP1 in particular, explained that performance tasks challenge her to think more deeply and she preferred performance tasks because the contexts relate to real-life situations.

LP1: The performance tasks are more challenging. We have to think harder. There are different and more ways of solving the problem and there are more answers... I prefer the questions because it relates to our everyday life.

Another student, AA1 voiced his preference for open-ended tasks because he believed the skills acquired through solving authentic and open-ended tasks could be transferred to other courses.

AA1: In the future, when we attend certain courses, we can use these types of skills for other courses….. For example, if I want to be a designer for a building, I could use these types of skills to help me design a building.

Challenges students' face in solving the tasks
Five items in part B of the post-questionnaire survey were designed to examine the challenges group E faced on the use of performance tasks. Table 8.19 shows the responses of group E students to the individual items.

Table 8.19. Distribution of response on challenges in solving performance tasks (Group E)

	Item (No of students =39)	DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q24	Doing mathematics performance tasks is difficult to me.	0.0%	2.6%	17.9%	5.1%	33.3%	10.3%	15.4%	0.0%	15.4%	5.54
Q28	I have to think harder when I am doing maths performance tasks	0.0%	0.0%	2.6%	5.1%	15.4%	20.5%	30.8%	10.3%	15.4%	6.64
Q29	I feel lost when I am doing maths performance tasks	2.6%	0.0%	12.8%	12.8%	25.6%	25.6%	7.7%	5.1%	7.7%	5.38
Q33	I need hints to help me do maths performance tasks	0.0%	0.0%	2.6%	10.3%	30.8%	20.5%	20.5%	7.7%	7.7%	6.00
Q35	Doing performance tasks takes more time than doing other maths questions usually done in class.	0.0%	5.3%	2.6%	2.6%	21.1%	2.6%	34.2%	13.2%	18.4%	6.61

About 46% of the students felt that working on performance task was difficult and approximately the same percentage felt lost when doing performance tasks. In the interview, LP2 reported a sense of frustration and felt discouraged when his attempts at solving the performance tasks failed.

LP2: Some of the questions require me to think very hard. I will try. If I can that is very good. But If I cannot, I call myself lousy and I will ask the teacher for help.

Researcher: Why do you think you are lousy?

LP2: Because I cannot do the performance task.

About three quarters of the students acknowledged the demands to ‘think harder’ in solving performance tasks and about 56% of the students expressed that they needed hints to solve the question. The teething problems faced as reported by LP1, LP2 and AA2 ranged from understanding the context of the question, the goals of the question to adopting an appropriate strategy to solve the problem.

> **LP1:** Sometimes it is difficult to get the first step. If you can overcome the first step to a question, then it is easier to get the rest of the question. if you cannot overcome the first step, then the teacher has to explain to you.
>
> **LP2:** When I come across the question for the first time, I usually don’t understand even though I read the question a number of times. I must ask the teacher what to do and the steps needed to solve the problem.
>
> **AA2:** Oh, sometimes I can’t make out what they want to find.

In a separate interview, Teacher E reaffirmed the range of difficulties faced by her students. She said: “*Most of the time they will actually ask me to explain what the questions are asking. … and how to do it?*

Though in the earlier part of the interview, some students had mentioned that Teacher E pitched her lessons at a level beyond them and they had difficulties understanding her explanation, she remained the dominant figure whom students sought for assistance if they were still unable to overcome their problems in the performance tasks. LP2 said: “*First thing I mean I ask the teacher. Then I try out what she had asked me. If I don’t understand, I will ask her again. From there also, I learn new methods, and from the new methods, if my friends don’t know, so I teach them. So we learn together”.* However, for the last 4 tasks, both LP1 and LP2 professed that they had found the performance tasks much easier.

Both high performing students and one of the average-ability students AA1, remained positive about working on performance tasks after the intervention period. They declared that they had no objections if performance tasks were integrated in the mathematics curriculum for the next semester.

> **AA1:** Manageable. Sometimes there are difficulties but I think we can manage with a little help.
>
> **Researcher:** If I were to carry on next year, would you mind having performance tasks?
>
> **HP1 & HP2 :** No.
>
> **Researcher:** You still want to do it despite all the challenges?
>
> **HP1 & HP2 :** Yes.

However, some students were undecided. In particular, AA2 expressed that performance tasks were too difficult and he needed to seek his friends’ help. Despite discussing with his friends, AA2 maintained that sometimes he still could not understand their explanation.

> **AA2:** I don’t really learn something new because it is too difficult. I have always asked my friends for help……Sometimes I can’t understand.”

Researcher: But when your friends helped you, do you benefit?

AA2: Not really ah.

However, in the later part of the interview, there was a change of view. All three students: two low performing students and the average ability student expressed that they did not mind having performance task at the end of each topic as a form of revision.

AA2: I think it is better to do it after every chapter.

LP1: Yes, it's like attempting a performance task after we have completed a chapter based on what we learned.

On the whole, all six students interviewed were positive about the use of performance tasks as an instructional tool in the mathematics classroom, in particular the two high performing students, HP1 and HP2, as well as the average ability student, AA1. One possible reason could be their sense of confidence with their own mathematical competence to solve authentic and/or open-ended performance tasks even though they had found the tasks challenging. For the other students, AA2, LP1 and LP2, they were apprehensive about continuing with the use of performance tasks but agreed to have performance tasks at the end of each chapter because they understood the benefits.

Students' perceived usefulness of performance tasks

In addition, 38.5% of the students perceived that doing performance tasks helped them to develop a more systematic approach to learning mathematics. This is in concurrence with one of the criteria in the rubric for grading performance tasks. With regard to the benefits, about 43.5% of the students believed they could learn mathematics by engaging in performance tasks. Student AA1 reaffirmed the same point in the interview when he said: *"I think these problem tasks are good for students, especially at our age we need to be exposed to different types of problem-solving".* Table 8.20 shows their responses to the 6 items designed to examine their beliefs on the usefulness of performance tasks.

Table 8.20. Distribution of response on usefulness of performance tasks (Group E)

	Item	DT	DAL	D	DLL	N	ALL	A	AAL	AT	Mean
Q25	Doing performance tasks helps me to learn maths.	5.1%	0.0%	12.8%	12.8%	25.6%	20.5%	17.9%	5.1%	0.0%	5.13
Q27	Doing maths performance tasks help me to be more creative in problem solving.	12.8%	0.0%	10.3%	7.7%	23.1%	20.5%	20.5%	5.1%	0.0%	4.97
Q31	Doing maths performance tasks helps me see more connections between maths and daily life.	5.1%	0.0%	7.7%	0.0%	30.8%	28.2%	10.3%	12.8%	5.1%	5.72
Q32	Doing maths performance tasks helps me to become more systematic when I am solving maths problems.	2.6%	0.0%	7.7%	7.7%	43.6%	10.3%	20.5%	5.1%	2.6%	5.44
Q37	Doing maths performance tasks makes me learn maths better.	7.7%	5.1%	7.7%	17.9%	41.0%	12.8%	2.6%	2.6%	2.6%	4.56
Q38	Doing maths performance tasks is a waste of time.	2.6%	2.6%	7.7%	10.3%	30.8%	7.7%	7.7%	7.7%	23.1%	5.95

About 56% of the students were positive that working on authentic and/or open-ended performance tasks helped them to perceive 'more connections between mathematics and daily life'. However, about 46% of the students regarded working on the tasks as 'a waste of time' and only approximately 23% disagreed with the statement.

In the interview with Teacher E, she expressed that students were concerned with the syllabus coverage for the semestral examination. Despite the knowledge that all performance tasks were designed to align with the syllabus, some students had misconstrued that performance tasks were carried out at the expense of the mathematics content and were concerned with the preparation for their semestral examination. She said: "they were asking why we are doing this? Why are we not doing the normal things? This is because they are quite concern with their work. They are scared that... this is at the expense of their syllabus."

The pragmatic mindsets of the students about academic achievement may have stemmed from many sources such as expectation of parents and the inherent structures present in the system, like edusave scholarships that reward individuals who excel. It seemed, to the researcher that the negative feelings towards performance tasks may be attributed to the mistaken belief that the interventions were implemented at the expense of the syllabus coverage which in turn may undermine their performance in the school-based examination.

Teacher E also reiterated the same point in a separate interview about the students' concerned with their academic achievement and would show greater enthusiasm if the scores in the performance tasks were reflected in the assessment score cards. She explained: "They will be interested because of the marks. They are quite a practical batch you know... They are certainly not intrinsically motivated."

To conclude this section, three quarters of the students in group E responded that working on performance tasks challenged them to think harder and encouraged them to adopt different approaches in solving the task. A quarter of the students in group E felt that they were good at solving mathematics problems and slightly less than half viewed performance tasks as beneficial towards the learning of mathematics. Teacher E remained the main person whom the students sought for assistance, though some students had expressed in the earlier part of the interview that her explanation was pitched beyond their learning ability and they would rather seek help from their friends. Despite the challenges that the two low performing students faced in the initial part of the intervention programme, they nonetheless found the last four tasks to be much easier than before. Less than half of the respondents in group E regarded working on the tasks as 'a waste of time'. The pragmatic viewpoints of the students towards academic achievement might have contributed to the negative feelings towards the tasks.

It was believed that exposure to authentic and/or open-ended performance tasks helped students to acquire problem-solving skills. This was also consistent with London's (1993) findings, who noted that non-routine problems which were not included in the traditional curriculum imbued students with higher-order problem-solving. In addition, he commented that students developed a sense of mathematical maturity when they laboured through a sequence of non-routine tasks. Hence, the ripple effect of integrating performance tasks might be felt through long term usage when the students were engaged in solving the tasks.

8.4.4 Results of the 'new strategy task' test

Taking into consideration that the students had just begun their secondary education, the tasks used in the pre-test only covered two main topics in the primary mathematics curriculum: arithmetic and geometry. Three open-ended mathematical tasks were specifically designed for the pre- test, with three parallel tasks for the post-test. Each of the tasks had more than one acceptable answer.

The pre-test consisted of one geometry task: '*Cube Arrangement'*, and two arithmetic tasks, '*Movie Selection'* and '*Filling Numbers'*, while the post-test consisted of 3 similar tasks: '*Cuboids Arrangement'*, '*Song Selection'* and '*Filling Numbers'*. To assess the students' performance on the open-ended tasks, task-specific rubrics were designed based on three criteria: Approach (A), Solution (S) and Presentation (P). A generic rubric[14] for grading the 'New Strategy Task' Tests is shown in Appendix 8.1**.**

The student's scores for the individual criterion: Approach (A), Solution (S) and Presentation (P) were totalled for the three questions. Two statistical tests were used in the analysis of the results for each criterion:

- the Mann Whitney U test was applied to verify the significant difference of the results for the pre- and post- 'new strategy' tasks tests between group E and group C.
- the Wilcoxon Signed Ranks test was used to confirm the significant difference between the pre- and post- 'new strategy tasks' tests within each group.

Table 8.21 indicates the results of the Wilcoxon Signed Rank test for the mean scores of A, S and P for the pre- and post-tests of groups E and C.

Table 8.21. Mean scores of Approach (A), Solution (S) and Presentation (P) in pre- and post-tests of groups E and C

	Group E				Group C			
	Pre-(40)	Post-(39)	z value	p	Pre-(40)	Post-(30)	z value	P
A	5.90	6.97	-2.941	**0.003***	4.58	5.97	-2.480	**0.013***
S	6.30	6.56	-1.291	0.197	3.33	4.67	-1.723	0.085
P	5.60	5.85	-1.060	0.289	4.75	5.07	-0.445	0.656

Note. *less than 0.05 level of significance for two-tailed test.

Comparing the mean value of A, S and P in Table 4.19 between the pre- and post-'new strategy task' tests for group E, there appeared to be a slight improvement in all the three criteria. Similarly for group C, a comparison between the pre- and post-test revealed an increase in the mean value of A, S and P. The increase appeared to be greater for group C as compared to group E.

Table 8.22 shows the results of the Mann Whitney U test between group E and group C for the pre-tests and post- tests of the 'new strategy task'.

As we can see from the table, in the pre-test, group E performed significantly better than group C for all the three criteria: A, S and P. However in the post-test, group E performed significantly better than group C only in the two criteria: A and S.

Table 8.22. Results of the Mann Whitney U Tests for Group E and Group C

	Pre-test					Post-test				
	E (40)	C(40)	U value	Z value	Sig	E(39)	C(30)	U value	Z value	Sig
A	5.90	4.58	342.5	-4.522	**0.000***	6.97	5.97	425.5	-1.959	**0.050***
S	6.30	3.33	263.0	-5.226	**0.000***	6.56	4.67	324.5	-3.186	**0.001***
P	5.60	4.75	535.0	-2.609	**0.009***	5.85	5.07	442.0	-1.757	0.079

Note. *less than 0.05 level of significance for two-tailed test.

As reported earlier, group E and group C were non-equivalent groups and therefore it would be difficult to infer the influence of the interventions implemented, by comparing the performance of group E and group C in the pre- and post-'new strategies task' tests.

In addition, the presence of other variables and constraints such as the absence of sample answer scripts for purpose of benching-marking in the scoring procedure of the pre- and post-tests may have also affected the results.

Both the pre- and post- 'new strategy task' tests were specifically designed for this study. Two pilot testing were carried out with the pre-test but no pilot testing was done for the post-test, based on the assumption that the post-test is parallel to the pre-test. A concern arose in the process of grading by the two raters. First, as sample answer scripts were not available for the purpose of bench-marking when the task-specified rubric was crafted, several points in the solutions and rubric were overlooked. After many discussions between the two raters, the rubric was amended with the inclusion of the points that were present in the students' answer scripts. The tasks and amended rubric for the post-test, was found in Appendix 8.5. The changes were highlighted and underlined in italic.

The answer scripts for the pre- and post-tests for group E and group C were then re-marked to ensure greater alignment of the two tests and the task-specific rubrics. At the same time, the task specified rubrics was further refined.

To conclude this section, group E appeared to have performed much better than group C based on the mean value of 'A', 'S' and 'P' for the three tasks with significant difference in the pre-test. However, in the post- test, group E only performed better than group C with significant difference for 'A' and 'S'. Besides the presence of the non-equivalent group C, the absence of sample task might influence the scores of the new strategy task tests, Therefore, the effects of the interventions on group E was not clearly visible using the new strategy task, pre- and post-tests. It seemed that Teacher E might need to give more guidance to group E on the presentation of solutions such as the organisation and documentation of the answers.

8.4.5 Results of students' performance on the interventions

A total of 10 interventions were carried out between April 2004 and May 2005. The last performance task which was completed on 15 May was graded but due to the Mid-Year Examination in May 2005, the answer scripts were not returned to the students before the month-long June holidays. When the academic term reopened in late June, Teacher E had resigned from the teaching service.

Response rate

The performance tasks were designed to be completed within the timeframe of two class periods. For the preliminary section, Teacher E was advised to go through the "warming up" exercises as class discussion for the first 15 to 20 minutes of the lesson, with the rest of the time allocated for the harder sections on the authentic and/or open-ended performance tasks. However, due to varying abilities of the students, the duration taken to complete the open-ended tasks differed. Some students completed the task within the stipulated timeframe of two periods and returned the answer scripts to the teacher. Those who were unable to complete the task would hand in their work on subsequent days as instructed by Teacher E.

All interventions carried out in the classroom were observed by the researcher and videotaped with the consent of Teacher E. The response rate for each task is indicated in Table 8.23.

Table 8.23. Response rates for individual tasks

Task No.									
1	2	3	4	5	6	7	8	9	10
River Cruise	What is the time now?	Painting a room	Physical Anthropology	Time Deposit	National Flag	Singapore Currency	Size does matter	Building Bricks	Decorate your Kitchen
35%	50%	100%	67.5%	70%	67.5%	60%	35%	87.5%	80%

The response rate varied greatly from 35% to 100% for the ten performance tasks as shown in Table 8.23. Some students who submitted their answer scripts for the tasks did not have their names on it, thus marking the process of record-keeping and monitoring very difficult. Hence, the researcher was unable to ascertain the names of students who consistently did not submit their answer scripts, which otherwise would be removed from the analysis of the data.

As reported earlier, the answer scripts were graded separately by the two raters, based on the pre-designed task-specific rubrics. A consensus on the performance levels for the three criteria was reached only after a discussion between the two researchers. No sample tasks were available for the purpose of benchmarking. Descriptive statistics was used to delineate the students' performance in the interventions for the three criteria: Approach (A), Solution (S) and Presentation (P).

Students' approach to intervention tasks
According to Pólya (1971), there are four stages involved in problem- solving: (1) understanding the problem, (2) devising a plan, (3) carrying out the plan, and (4) looking back. The criteria on "Approach" as spelled out in the rubric referred to the first two stages of Pólya's four stages of problem solving.

In most of the interventions, the strategy adopted by the students was based on trial and error. Very few students explored beyond one approach or adopted a systematic approach to solve the tasks. With reference to the ten completed tasks, the number of students who managed to obtain a Level 3 for the criteria on "Approach" was as follows:

- only one student obtained a Level 3 for the task on 'Time Deposit'
- five students obtained a Level '3' for the task on 'Singapore Currency'

The rest of the students only managed to attain a Level 2 or less for all the tasks. Based on that performance level, it appeared that the strategy commonly adopted by students to solve the tasks was 'trial and error'.

The strategy of 'trial and error' may be effective at the initial stages of problem-solving and for problems with relatively simple situations. However, it is certainly not an efficient strategy for open-ended tasks because it may not lead to a complete solution as the number of trials can be extensive and time-consuming. Hence it appeared that the enhancement of problem-solving skills is necessary to expose students to other approaches besides 'trial and error'.

In an interview carried out after the intervention period, one average- ability student disclosed that he frequently used the strategy of trial and error because he was not clear in his understanding of certain concepts for topics such as Algebra.

> **Researcher:** When it comes to solving this task such as Decorating of Kitchen, or Laying of tiles, what sort of approach do you use?
>
> **AA1:** I normally used trial and error…. because I am weak in Algebra.

There appeared to be a need to review certain learning difficulties experienced by students and mathematical concepts for some topics, such as algebra. The researcher was unable to relate the information to Teacher E because she had already left the teaching service after all the interventions had been carried out.

'Openness' of students' solutions
To solve the authentic and/or open-ended performance tasks, students needed to take into account the context of the question and the different conditions present in the question. In the interventions graded by the two members of the MAP research team as raters, it seemed that most students had only considered one condition that was necessary to produce a single answer. The presence of different conditions which might lead to multiple answers in the tasks was overlooked. With the exception of the two tasks, Task 2 and Tasks 9, most students did not produce more than one answer. None of the students obtained a level 4 for the criteria on "solutions" in any of the tasks. Table 8.24 shows the percentage of students who produced more than one solution for the various performance tasks.

Table 8.24. Percentage with more than one solution for tasks

Tasks	1(i)	(ii)	2	3	4	5	6	7	8	9	10
More than[1] one solution	7.7%	0%	91.3%	NA	0%	20%	5.5%[2]	NA	6.5%	80%	40.6%

Note. [1]The percentage was calculated based on the number who submitted their answer scripts. [2]A general formula is required in this case. Though the context of the question is authentic, the task does not require multiple solutions or it is too complex to consider more than one solution.

It was not clear if the underlying problem was the students' inherent belief or habit that it sufficed to produce just one answer to the mathematics problem, though they articulated in the interviews that the performance tasks contained multiple approaches and answers. However, it was likely that students might need more exposure to open-ended performance tasks to become accustom to coming up with more than one acceptable answer. It was also noted that the number of students who considered more that one solution increased substantially for the last two tasks.

Presentation of students' solutions
The rubric of the performance tasks includes presentation of the solutions as one important aspect of students' performance in problem-solving. Five levels are prescribed in the rubric, ranging from Level 0 – 'No attempt is evident to Level 4 – 'Presentation of work is clear and greatly adds to the markers' understanding of the working procedure'.

In the analysis, the percentage of students who obtained a performance Level 3 for clarity and systematic documentation for all the interventions is shown in Table 8.25.

Table 8.25. Percentage with a Level 3 for the aspect on presentation

Tasks	1(i)	(ii)	2	3(i)	3(ii)	4	5	6	7	8	9	10
Percentage with a Level 3* on P	28.6%	21.4%	0%	2.5%	2.5%	0%	39.3%	5.6%	3.8%	3.0%	2.9%	12.5%

Note. * The percentage was calculated based the number who submitted their answer scripts

The data collected revealed that documentation in most solutions was not organised and many solutions contained only computations or procedural workings. With the exceptions of Task 1, Task 5 and Task 10, the percentage of students who obtained a Level 3 for documentation was less that 10%, as shown in Table 8.25. Even the researcher, who crafted the questions and rated the performance tasks, encountered much difficulty in understanding the students' solutions because of missing statements or disorganised documentation. Hence students in group E might need guidance from their mathematics teacher in the presentation of solutions.

In the interview with two low-performing students, they reflected that the last four performance tasks appeared to be much easier than those implemented at the earlier stage. One possible explanation could be that students needed more opportunities to become familiar with the "new strategy" tasks, the context as well as the expectation of the answers before being comfortable at solving these problems.

To conclude this section, the findings from the implementation of the ten interventions seemed to suggest that the use of authentic and/or open-ended performance tasks have, to a varying degree, a positive influence on the students.

8.4.6 Results of the interview with Teacher E

In this study, Teacher E who managed to integrate ten performance tasks within a timeframe of eighteen months spoke very positively about the use of performance tasks in her mathematics lesson and was convinced of the benefits. She expressed that their use not only promoted students' interest in the learning of mathematics but also engaged them in the process of solving non-routine problems which were different from the textbook problems. She said: "*The performance tasks will create interest in them... expose them to non-routine questions that are different from the textbook*."

When queries were raised about the difference between the routine problems found in textbooks and the authentic and/or open-ended performance tasks, Teacher E articulated that: "*The regular exercises are geared towards drill and practice while the open-ended tasks focus more on applications on what they have learned."* In addition, she explained that routine homework usually necessitates students to work out the answers. However, working on performance tasks required students not only to focus their attention on obtaining the answer, but also provide the opportunities for them to explore further and draw conclusions. Here is an extract of the interview, she said: "Most of *the math questions which the students had attempted usually do not require a conclusion only the solution. In the case of the performance tasks, the students not only have to solve the question, they have to draw some conclusions. It's like doing an experiment.*"

Her positive attitude is one of the critical factors that promote the use of authentic and/or or open-ended questions in her class. There were a number of informal discussion sessions which Teacher E had engaged with the researcher, so that she could effectively integrate the tasks in her lessons. Her enthusiasm and her sense of responsibility to plan her delivery so that she could better guide the students when they were attempting the tasks were commendable. In fact Teacher E who was so convinced of the benefits of using the performance tasks, had expressed her willingness to replace 20% of the regular homework with the "new strategy" tasks. She said: "*Around twenty or thirty, I think is fine."* All interventions carried out with group E were observed by the researcher.

In the months of intervention, Teacher E observed and was pleased that the students persisted in their efforts despite the challenges faced in obtaining the answers. *She said:"…they tried even though they were not successful. Maybe one or two may have done some research."* Besides the difficulties of understanding the context of the questions and adopting the right approach, Teacher E also voiced that the students were unclear about the goals of the question. She said: "*Most of the time they will actually ask me to explain what the questions are asking... and how to do it?*

Teacher E also commented that the strong camaraderie among the students among her students might have provided the much needed support to overcome the difficulties in solving the tasks. In the interview, she said: "*Even though there are times when they do not understand they ask one another. They learn from their friends and they are certainly not the weaker ones. They are quite bright.*

The difference between the grading of daily homework and the performance task was clearly articulated by Teacher E in the interview, though she largely did not participate in the marking of the ten interventions, with the exception of five selected randomly scripts on one occasion. She commented that grading with the rubric was more encompassing as it contained more details in comparison to her marking of daily homework, which is restricted to just a right or wrong answer. She said: "*Definitely, this is more detailed, I mean you can assess different levels. If I mark the normal daily work, it will be based a lot on right or wrong."*

Teacher E clarified that her initial reluctance to partake in the grading of the interventions was because she was not familiar with the different levels in the rubric. She explained: *"initially, I'm not quite sure about like how we really put the level*." However if the rubrics were provided, Teacher E gave the assurance that she does not mind grading similar tasks. She said: "*Similar tasks, maybe I can try*."

Through the interview, it seemed that the enthusiasm and positive attitude of Teacher E had contributed significantly to the success of this study. The impact on the students would be greater if she had participated in the scoring of the tasks. In that way, Teacher E would not only be able to effectively infuse the tasks in her lessons but also provide feedback to students on their shortcomings and guide them to make improvements in subsequent tasks. However, Teacher E left the teaching service after all interventions were implemented.

8.4.7 Summary of results

To conclude this section, the findings in this sub-study seemed to suggest that the use of performance tasks as an alternative assessment strategy has positive effects on the academic achievement of students in the school based examinations. The mean difference between group E and group C was raised substantially from 8.4 for F(2004) to 12.0 M(2005). However, due to the presence of a non-equivalent comparison group, the extent of the effects cannot be clearly determined.

The ten interventions carried out over a period of eighteen months appeared to have, to a certain degree, a positive influence on the affective component of the students. Though the affects of both groups appeared to have declined over the eighteen months of interventions, this phenomenon is not an uncommon occurrence among growing-up school-going teenagers. Group E appeared to remain more confident of their mathematical ability and less anxious to face challenging problems as compared to the comparison group. However, a more in-dept study needed to be carried out to determine the extent of the impact on the students.

The interventions carried out over the past eighteen months appeared to have a positive influence on some students. A good number of the students in group E responded that working on performance tasks challenged them to think harder and encouraged them to adopt different approaches to solving a task. A quarter of the students in group E even felt that they were good at solving mathematics problems. It seemed that the students needed to work on a number of performance tasks before they could be comfortable at solving performance tasks. This was reaffirmed in the interview where two low performing students articulated that they were able to manage the last four tasks as compared to those interventions carried out at the initial stage of the intervention programme.

Some students in group E regarded working on the tasks as '*a waste of time*', while others viewed performance tasks as beneficial towards the learning of mathematics. The negative feelings towards performance tasks may have been attributed to the pragmatic viewpoints of the students towards academic achievement and the mistaken notion that performance tasks were integrated at the expense of syllabus coverage.

Teacher E's enthusiasm and her positive attitude were crucial in promoting the use of authentic and/or open-ended questions in her class and the success of this study. There were a number of informal discussion sessions which Teacher E had with the researcher, so that she could effectively integrate the tasks in her lessons. Though the students had commented that Teacher E was overly strict and her explanations were frequently pitched at a level beyond their learning ability, she remained the main figure whom some students sought for assistance in the face of difficulties.

On the whole, the interventions seemed to have a positive impact on Teacher E. She was assured of the benefits and was willing to replace 20% to 30% of her usual homework with performance tasks. In fact she appeared to feel pleased that her students had not only persevered but strongly supported one and another in face of challenges when solving the open-ended tasks.

The students' performance on the ten interventions seemed to suggest that more exposure to authentic and/or open-ended tasks was necessary. It appeared that the strategy commonly adopted by the students to solve the authentic and/or open-ended performance tasks was based on 'trial and error', an inefficient strategy. The reason behind the adoption of such an approach might be due to the learning difficulties that students faced in understanding certain concepts for topics such as Algebra. Hence, it appeared that the enhancement of problem-solving skills and the review of difficult concepts were necessary to expose students to other approaches, besides 'trial and error'. In addition, most students produced only one answer for the open-ended tasks. It was not clear if the underlying problem was due to the students' inherent belief or habit of the mind that it suffices to produce just one answer. However as the number of students who considered more that one solution increased substantially in the last two tasks, it might seem that students required more exposure to the performance tasks to be comfortable in producing more than one acceptable.

The percentage of students who obtained a performance Level 3 for clarity and systematic documentation for three interventions carried out was less than 10%. The rest of the students obtained only a lower performance level for the criteria on presentation for all the tasks. Hence, guidance is certainly necessary in the presentation of solutions for some students who lacked the ability to organise and present their solutions in a systemic manner.

On the whole, the authentic and/or open-ended performance tasks seemed to have a positive influence on the cognitive and affective domains of the students as well as on the pedagogical practice of the teacher. The interventions provided the teacher with opportunities to expose and engage students in solving performance tasks with multiple strategies and multiple solutions. Besides improving problem-solving skills, the authentic context of the tasks helped students to relate mathematics to real-life situations. In addition, the non-routine and open-ended nature of the problems motivated the students to persevere in face of challenges and fostered the positive perception that they were able to solve difficult problems.

8.5 Summary, Conclusions, and Recommendations

8.5.1 Summary and conclusions

This sub-study investigated the effects of integrating authentic and/or open-ended performance tasks in classroom instruction on students' learning of mathematics in a non-high-performing or typical neighbourhood secondary school. The study yielded several findings which provided answers to the four research questions.

For the first research question, which is about the effects of integrating performance tasks in classroom instruction on students' mathematical achievement in school-based examinations, it was noted that the difference between the mean score of group E and group C for M(2004) was approximately 9.7. However the difference was reduced to 8.4 for F(2004) and widened to 12.0 in M(2005). According to the interviews with the students and school administrator, it seemed that the difference of 12 points in the semestral scores of M(2005) between group E and group C, -an increase from 8.4 points in F(2004), might not be attributed largely to the positive influence of Teacher E. On the contrary, it was the mathematics teacher of group C who appeared to relate well with the comparison class. Hence the gap between the mean scores of group E and C, in favour of group E seemed to suggest the positive influence of the interventions on the academic performance of the students in group E. However, caution needs to be taken in interpreting the results as the design of the study includes a non-equivalent comparison group.

For the second research question which examined the effects of integrating performance tasks (authentic and/or open-ended) in the mathematics classroom on students' affects towards the learning of mathematics, the pre- and post-questionnaire survey was used. In this study, the open-ended performance tasks were defined as non-routine problems where the solutions or goal was not evident and there was no obvious algorithm or procedure for students to use. Hence when students engaged in solving the open-ended tasks, they might experience an array of feelings. Students who failed to reach a satisfactory solution might report feelings of frustration as indicated in the interview with LP2.

This section presents a brief summary on affective issues in relation to the use of authentic and/or open-ended performance tasks. Students' affective reactions to solving performance tasks are influenced by various factors such as their personal perceptions, perceived competence, feelings of anxiety and the perceived usefulness of learning Mathematics. For example, decisions about whether to persevere along a solution path may be to some extent, influenced by the student's sense of confidence or level of anxiety.

The analysis of the pre- and post-questionnaire surveys administered to group E and group C was presented in accordance with the four domains: general perception, anxiety, perceived competence and perception on usefulness of mathematics. On the whole, the findings indicated that classroom interventions carried out over a timeframe of about eighteen months appeared to have a positive influence on the affects of the students.

Under the domain of general perception towards mathematics, the large gap of 30% between the responses of group E to group C for the post- survey on their perception that '*Mathematics is hard for me'* could be attributed to the exposure of group E to the performance tasks. Their participation in this study might have helped the students in group E to acquire problem-solving skills that were helpful to handling non-routine and open-ended problems, thereby reinforcing their positive perception that they were able to cope with difficult mathematics problems. The findings were reaffirmed in the interview with two low performing students who revealed that they were able to cope with the later few performance tasks.

On the whole, in the pre- and post-survey, group E appeared to feel less anxious as compared to group C. As mathematics anxiety is inversely related to self-confidence, group E was more self-assured towards the learning of mathematics as compared to group C, which might be ascribed to the interventions in this study. Students developed self-confidence when they were provided with opportunities to engage in open-ended tasks as they were more challenging and demanding in comparison to routine textbook questions. The classroom presentation and discussion might have also helped to reinforce the positive affects when students communicated their solutions to the entire class.

An average-ability student, who was interviewed, revealed that performance tasks, though challenging, were nevertheless manageable when help was available. However, he felt apprehensive about solving more performance tasks in the forthcoming semester because he felt that there were too many things to learn. It appeared that the average ability student after the intervention programme, somehow came to realise that he needed to improve his content knowledge and develop his skills in several areas. Two high performing students who had been achieving in the traditional semestral examinations and had high expectations of themselves also shared similar feelings of apprehension. Both students were concerned that they would not be able to peak in their performance or they might lack certain mathematical conceptual knowledge when performance tasks were integrated to the mathematics lessons.

On the whole, comparing the pre- and post-questionnaire survey it appeared that the affects of group E and group C declined though they remained positive towards the learning of mathematics. The literature review on attitudes of students suggests that the decline is a natural process of growing up, especially when they progress from primary to secondary education. The analysis of the pre- and post-survey reflected a similar trend for both groups.

Although it was widely claimed that the student self-confidence declined as they progressed from elementary levels, the findings seemed to suggest that group E remained confident towards learning mathematics and solving mathematics problems, after the intervention period as compared to the group C. This is also evident in the number of question items that were significant in the post-questionnaire survey between group E and group C, in favour of group E:

- Mathematics is hard for me,
- It makes me nervous to even think about having to do a maths problem,
- I have a lot of confidence when it comes to maths,
- I can get good grades in mathematics.

The results seemed to suggest implicitly the positive influence of the new assessment strategy on students' attitudes towards mathematics. However the extent of the changes that can be attributed to the use of the new assessment strategy needs a more in-depth study.

The third research question investigates the effects of integrating performance tasks in classroom instruction on students' problem-solving abilities for open-ended questions.

The pre-test consisting of three mathematics tasks: 'Cube Arrangement', 'Movie Selection' and 'Filling of Numbers', was administered before the intervention. Similarly, the post-test consisting of three parallel tasks: 'Cuboids arrangement', 'Song Selection' and ‘Filling of Numbers’, was administered after the intervention. Task-specific rubrics were designed based on three criteria: Approach, Solution and Presentation.

Comparing group E and group C, the results indicate significant difference for all the three criteria, A, S and P, in the pre-test but only significant difference for the two criteria, A and S in the post-test. However the mean value of A, S and P for group E is higher than group C which seems to suggest that group E performed better than group C for the pre- and post-test.

The response rate for the ten interventions carried out varied between 30% and 100%. However due to the missing names on some answer scripts, it was not possible to consider only those students who had consistently submitted their answer scripts. It was not clear what the underlying reason for not submitting the scripts were but the interview with Teacher E appeared to suggest that the students’ mistaken notion of performance tasks being carried out at the expense of the syllabus coverage thereby affecting their grades for the impending examination as one of the causes.

The task specific rubric was developed to grade the performance tasks based on three different criteria: A, S and P. For the criterion on “Approach”, most of the students relied primarily on a trial-and-error strategy for most tasks but a handful of students adopted a systematic approach. Similarly, most solutions produced by the students satisfied only one of the two or more conditions in the problem. None of the students attained a level 4 for solution - the requirement for the attainment of a complete solution or a generalized formula. It was not clear if the underlying problem was the students’ inherent belief that it sufficed to produce just one answer to the mathematics problem, though they articulated in the interviews that the tasks contained multiple approaches and answers. However for the last two interventions, the number of students who produced more than one solution increased to 80% for performance tasks 9 (response rate is 87%) and 40.6% for performance task 10 (response rate is 80%). The findings seemed to suggest that students needed more exposure to the authentic and/or open-ended performance tasks to acquire the habit of producing more than one solution.

The documentation in most answers scripts contained only computations or procedural workings. The mathematical reasoning or explanation for the underlying processes was absent in most tasks. Peressini and Webb (1999) noted that one of the challenges faced by both students and teachers when open-ended tasks were incorporated into a mathematics classroom was the need for students to realise the importance of communicating their mathematical reasoning clearly, which could be presented in a variety of formats. Without proper documentation of the solutions, the researcher found it difficult to determine if the processes were in place or that the student actually knew and could solve the problem.

The fourth research question examines the influence of the use of performance tasks on teachers and students in the daily teaching and learning of mathematics.

1. Influence on Students in the Learning of Mathematics

A quarter of the students in group E responded that they were good at solving performance tasks and the same percentage liked the feature on multiple solutions. More than half responded that they enjoyed working on questions with multiple approaches and about the same number felt that they liked mathematics problems that relate to real-world situations. About three quarters of the students acknowledged the demands to ‘think harder’ when solving performance tasks.

The results were reaffirmed in the interviews with six students where all six had found the performance tasks more challenging and demanding as compared to routine problems in the textbooks, which they deemed only required the application of procedural knowledge.

Students also acknowledged that the open-nature of the problem demanded them to spend more time thinking about the different approaches and situations as they worked out multiple solutions. One student articulated that she preferred performance tasks because it related to real-life situations and another believed that he could transfer his skills to other courses.

The teething problems experienced by the students ranged from understanding the context and the goals of the question to adopting an appropriate strategy to solve the problem. Teacher E remained the dominant figure that some students sought for assistance, though they expressed that her explanation was pitched at a much higher level than their learning ability and would rather sought clarification with their peers. Interestingly, both the low-performing students professed in the interview that they had found the last four performance tasks to be much easier. This seems to suggest that students needed more opportunities to become familiar with the “new strategy” tasks. Hence more exposure to authentic open-ended questions would help to familiarise students with the context as well as expectation of the answers.

About 50% believed that engaging in performance tasks helped in the learning of mathematics and about the same number perceived that doing performance tasks helped them to make *‘more connections between mathematics and daily life*’. However, 46.2% of the students regarded working on the tasks as ‘*a waste of time*’. The misleading notion held by the students that the interventions were implemented at the expense of syllabus coverage might have contributed to the negative feelings towards the tasks.

On the whole, it seems that the ten performance tasks integrated in the mathematics instructional practice over a period of eighteen months might have had a positive influence on the students. At the same time, the integration of the ten performance tasks in the mathematics lessons seems to have helped students acquire problem solving skills.

2. Influence on pedagogical practice

Teacher E who managed to integrate ten performance tasks within a time frame of eighteen months was very positive about the use of performance tasks in her mathematics lesson. She was convinced that the exposure of authentic and/or open-ended performance tasks helped students not only in the learning of mathematics but also engaged them in the process of solving non-routine problems which are different from textbook problems.

Her positive attitude is a critical factor in promoting the use of authentic and/or or open-ended questions in her class. There were a number of informal discussion sessions which Teacher E had engaged with the researcher to discuss the worked solutions of the performance tasks prior to their implementation, so that she could effectively integrate the tasks in her lessons. All interventions carried out with group E were observed by the researcher.

Besides the difficulties of understanding the context of the questions and of adopting the right approach, Teacher E also voiced that students were unclear about the goals of the question. Nonetheless, she had tried her best to guide them during the implementation of the interventions. When carrying out the interventions in the class, Teacher E observed that the students persisted in their efforts despite the challenges faced. She also commended the strong camaraderie among the students – which helped to provide the much needed support that was necessary to help them overcome the difficulties faced in solving the tasks.

Teacher E, who was convinced of the benefits in using authentic and/or open-ended performance tasks, had expressed her willingness to replace 20% of the regular homework with the “new strategy” tasks. Through the interview, it seemed that teacher E was very positive about performance tasks and the use of the tasks in her instructional practice. However, she left the teaching service when the study was completed.

8.5.2 Limitations of the study

Like many other studies, there are certain limitations of the sub-study that may have implications on the interpretation of the data.

First, Due to the presence of a non-equivalent comparison group, as both the experimental and comparison class were nominated by the school administrator, the interpretation of the data may be affected. It would certainly be ideal if the comparison group was equivalent to the experimental group. Nevertheless, careful considerations have been taken in the interpretation of the data in this dissertation (see Chapter 3 and 4 for more details)

Second, the performance tasks were designed to align with the topics listed in the Scheme of Work and these tasks were sent to Teacher E via email. The selection and administration of the tasks were dependent on factors such as the suitability of the tasks to meet the needs of the students as deemed by the teacher and her teaching schedule for group E. The researcher would discuss with Teacher E the solutions and how she could effectively implement them in the class but did not impose her views. Hence the interval between the implementation of the different tasks varied. For example between Tasks 7 and Tasks 8 the interval was approximately three months but between Tasks 8 and Tasks 9 the interval was only 2 weeks.

Third, though teacher E was very positive about her involvement in the study, her reserved and stern manner in managing the class was somehow misconstrued by her students as being aloof and unapproachable. This was confirmed in separate interviews with the students and the school administrator. It would be desirable if the students had perceived her to be otherwise.

It is essential that the teacher involved receives training on the new assessment strategy and be adequately prepared for the implementation of the interventions. However due to her school commitments and tight teaching schedule, Teacher E was unable to attend the sharing sessions conducted by the MAP team, on 7 February 2004, prior to the implementation of the interventions, and on 6 November 2004 to update teachers on the progress of the interventions. The third sharing session was held on 15 July 2006, a year after the intervention period had ended - with the intent to brief teachers on the outcomes of the study but Teacher E had already resigned from the teaching service.

At the onset of the study, as Teacher E was unable to attend the session sharing on 7 Feb 2004, a separate meeting was arranged prior to the implementation of the interventions to share with her the rationale for this study as well as the features of “new strategy” tasks: open-endedness and authenticity. However, in the interview conducted immediately after all the interventions had been carried out, Teacher E recounted that when the students enquired about the class involvement in the use of “new strategy” tasks, she gave the answer that it was a random choice. She said: “*Yes. They ask and I tell them they are selected by default.”* Hence it appeared that rationale of the study was not appropriately conveyed to the students, resulting in presumptions that the tasks were carried out at the expense of the syllabus coverage.

It was the researcher’s intent to provide Teacher E with the performance tasks at the start of this study, so that she could adapt them to match the learning abilities of her class and in the process, learn the techniques of crafting performance tasks and independently proceed to design the tasks for her class. However due to her heavy teaching duties, Teacher E did not adapt the pre-designed tasks, nor did she attempt to design any tasks for her class. In the interview after the intervention programme, she explained that besides the time constraint, it was beyond her ability to craft any task or rubric.

A sharing session on the grading of the tasks with the use of the rubrics was arranged in December 2004, after the implementation of six interventions. Teacher E had randomly selected five scripts from some tasks to grade them on her own. A follow-up discussion on the grading process using rubrics was carried out with the researcher. However, Teacher E was not involved in the scoring of the tasks in the study, though the researcher had shared with her, the skills in scoring the tasks using the rubrics. It would be ideal if Teacher E had been involved in the scoring process. In that way, she would be able to provide feedback to the students on their shortcomings and monitor their progress for subsequent interventions.

Finally, the researcher was aware that the solution of the performance tasks which the class had attempted was put up on the class notice board. Though the solutions and scores with comments were returned, not all students attempted to find out their mistakes. Only one (LP1) out of six students interviewed remarked that she had scrutinised the comments, answer scripts and scores, and also made comparisons with the given solution. She commented: "*When I got the paper, I looked through the comments first and the grades. After that I looked through the solution and compared it with the given solution.*"

Similarly, neither did teacher E work through the solutions with the class nor provide any comments on the answer scripts when they were returned to the class. She explained that she was too engaged in her lessons and had persistently forgotten about them. The researcher felt that it would be more fruitful if the teacher had set aside fifteen minutes of her lesson time to return the graded answer scripts and go through the solutions with her students. However, the tight work schedule and the unspoken need to complete the syllabus to prepare students for the impending examinations, made it difficult for Teacher E to do so especially after she had spent two periods or more on the performance tasks.

8.5.3 Recommendations for further studies

This study focused on the effects of authentic and/or open-ended performance tasks on the academic achievement of students in school- based examinations, the affective aspects of the students, their problem- solving abilities and on the teaching and learning of mathematics. The subjects of this study were Secondary One students from a neighbourhood school. The researcher believes that these tasks are also applicable to Secondary One students of a high performing school. It would be interesting if the study was replicated on students of different age groups in primary schools. One recommendation for further studies would be for the extension of the study to include primary pupils of different levels.

The study has also suggested that teachers play a crucial role in the implementation of the new assessment strategy. It would be beneficial if the study was extended to investigate the teacher-related variables that help to promote its use in the mathematics classroom such as:

- the teachers' knowledge of mathematics and the affective aspects in her personal experience;
- the teachers' beliefs about the teaching and learning of mathematics and how these relate to the implementation of alternative assessment. For example, a teacher who believes that her primary mission is to teach mathematics so as to prepare students for the impending examinations, may not be enthusiastic in the way she implements the tasks;

- the teachers' experience in the classroom, such as the length of service and even the amount of exposure to the new assessment strategy prior to the study, may influence their actions in the integration of such tasks in their lessons.

Further studies on the 'teacher related factors' may help to shed light on how authentic and /or open-ended performance tasks can be effectively implemented in the mathematics classroom.

Chapter 9 Results and Findings (VII): Communication Tasks (Secondary)[15]

9.1 Introduction

This chapter reports the results from the sub-study of the MAP project focusing on using communication tasks as the alternative assessment strategies in two secondary schools.

Developing students' communication skill is believed to be one of the important goals in mathematics education. Singapore's mathematics syllabus emphasizes that students' ability 'to communicate mathematically… to be able to illustrate, to interpret, to explain and to discuss mathematical ideas and experiences in doing mathematics' (MOE, 2000, p. 17) is an important communication aspect in learning mathematics. Students should be given opportunities to speak and write substantially even in mathematics classrooms although very often mathematics learning has been viewed to focus mainly on computational and procedural skills.

In this sub-study, communication tasks were classified into two types: written tasks and oral tasks. More specifically, this study had employed the use of journal writing tasks as written tasks and oral presentation tasks as oral tasks. Although the use of journal writing and oral presentation is not prevalent in the context of our local schools, some small-scale structured research works in the mathematics classrooms (e.g. Yeo, 2001; Yazilah & Fan, 2002; Seto, 2002) seemed to suggest the increasing recognition and interest in the use of these alternative strategies. However, research documentation of how mathematics teachers can efficiently engaging students in meaningful learning by using journal writing and oral presentation in secondary mathematics classrooms is under-researched (also see Hadjioannou, 2007).

The study was intended to investigate the effects of using journal writing and oral presentation in Singapore secondary school students' daily mathematics learning, and explore possible ways to help teachers integrate these new strategies effectively and efficiently in their daily teaching practices.

9.2 Research Questions and Conceptual Framework

9.2.1 Research questions

As said earlier, the large research project has two main research questions focusing on the effects of using new assessment strategies on students' mathematics achievement in both cognitive and affective domains, and how teachers can effectively and efficiently integrate new assessment strategies into their daily mathematics classroom teaching instructions.

Likewise, this study has the following specific research questions:

(1) What are the effects of using journal writing and oral presentation tasks in daily secondary mathematics classroom teaching on students' attitudes towards mathematics and learning of mathematics?
(2) What are the effects of using journal writing and oral presentation tasks in daily secondary mathematics classroom teaching on students' school mathematics achievement?
(3) What are the effective ways of implementing journal writing and oral presentation tasks in the daily mathematics classroom learning?

9.2.2 Conceptual framework

Journal writing tasks

Although different research studies have different ideas of what journal writing is but there seems to be a consensus that a journal writing task must be a piece of work that student writes through which teacher can obtain useful information about mathematics learning and teaching and then make inferences for a variety of purposes (Robinson, 1998). In this study, journal writing is a task whereby student is given the opportunity to write through to reveal his or her thinking, reflection, understanding and feeling concerning mathematics, mathematics learning, and mathematics teaching. While there is no one or formal format of what exactly must be in the journal writing, the most important point for this study was that journal writing must be embedded in the structure of the mathematics curriculum and therefore journal writing should be integrated in the mathematics instruction. In this study, journal writing tasks are classified into four different categories depending on the purpose for the writing:

- *Content* tasks are designed to allow students to recall mathematical knowledge by defining, making a summary, to state a concept/term, to form an over view or give examples of what have been learnt earlier;

- *Process* tasks are designated to allow students to demonstrate their mathematical explanation and reasoning of working on a solution or idea;

- *Reflection* tasks are intended to provide students opportunity to do self-reflection on their own learning of what has been learnt and give thoughts to how they could improve their own learning;

- Task in *Affective domain* focuses on writing that allows students to express their feelings, interests, beliefs, perseverance and attitudes toward their own learning and their teachers' teaching.

Samples of the journal writing tasks based on the different categories are shown in Table 9.1.

Table 9.1. Samples of journal writing tasks

Content task	• Today I am going to explain to you the differences and similarities between *Percentages, Fractions and Ratios* • You have learnt the concept of prime numbers. Now, suppose you are a teacher and Jason is a new student who has not leant the concept of prime numbers. Please describe in detail how you would explain to Jason what a prime number is.
Process task	• Suppose your good friend, Joe was absent when the teacher was teaching 'use of brackets in simplification'. Can you explain to Joe how to simplify: 5a +2d – 2 {3a – 3 – (2b – c + b)} – 4c? • Is $2a^4$ the same as $(2a)^4$? Explain your answer.
Reflection task	• I am thinking of a particular problem/difficulty I had during the exams. Now I know why I made mistake and didn't understand earlier. Let me explain to you what I didn't understand earlier but now I learnt….. • The mathematics that I learnt in the last two weeks are:… However I still find it difficult to work on:…. Reasons being:…..
Affective task	• I would like to tell my teacher how I feel about asking me to write and explain my understanding of mathematics. Complete the letter to your teacher! *Dear Teacher,* *This is how I feel……* • My aspirations! *My target for my math grade is:….* *In order to achieve my target, these are what I should do and my challenges are:….* *My concerns and worries are:…..And this is how I am going to deal with them:….*

Oral presentation tasks

While there is no one formal definition of what is oral presentation, it is evident in the literature that oral presentation is an activity of sharing ideas and clarifying understanding verbally. The purpose of oral presentation is to allow teacher to hear what students are thinking about mathematics, and how they express mathematics and the understanding of mathematics in their own words. This study adopts the ideas from the *Communication Standard for Grades 6-8* (NCTM, 2000, p. 272), that oral presentation task must provide opportunity for students to think through questions and problems; express their ideas and reasons; demonstrate and explain what they have learnt; justify their own opinion; and reflect on their own understanding and on the ideas of others. The above points provided practical suggestions for designing and implementing oral tasks in daily teaching and learning.

In this study, oral presentation tasks are meant to be integrated into the daily teaching and therefore the topics of presentation should be chosen that support the current instructional content and syllabus. The framework that is developed in this study to represent specific oral presentation tasks is generally classified as pre-structured, designed or prepared and tasks that are not pre-planned but impromptu (i.e. tasks are carried out without being planned earlier or rehearsed) depending on the criteria or purposes. Table 9.2 gives a brief description of what impromptu and pre-structured oral presentation tasks are

Table 9.2. Description of impromptu and pre-structured oral presentation tasks

Impromptu tasks	Pre-structured tasks
Specific tasks: • Questions that are posed during instruction; • Students' responses to questions that are posed or asked; • Students' work that are represented on the board during instruction; • Students' work that are given in the homework, class • work, worksheets, or textbooks; • Students' summary of each day's lesson.	Specific tasks: • Questions that are pre-designed and given to students prior to lessons; • Students' previous writing tasks on their learning reflection or perceptions; • Students' previous solutions to test questions; • Students' previous writing tasks about mathematics; • A selected topic that is pre-agreed before discussion; • To report results or findings of a project work; • To report pair or group work discussion.

To illustrate how a pre-structured oral presentation task could look like, take for example in the topic of Hire Purchase, teacher could ask students to bring in newspaper cuttings and advertisement about purchase of items that involved both cash and hire purchase scheme. The questions that a teacher wants students to think and work on before the presentation may look like this:

(a) Discuss and write down that is agreed among yourselves in the group the understanding about hire purchase.

(b) Find an advertisement which has both cash and hire purchase scheme. Cut out and paste the advertisement on the worksheet.

(c) Suppose you are to advice your parents about purchasing the product (that you have cut out earlier in part (b)), what would you say to them and what advice would you give?

In another example, a teacher may want to collate the common mistakes that students have made in a test or examination about the expansion or factorization in algebraic expressions. The worksheet should contain the work examples of students that have made mistakes in their test/examination, a column for students to work out the correct solutions, and a column for students to write out how mistakes pertaining to the specific questions could be prevented. After some time of either individual or small group discussion, students can therefore present and discuss about their work and ideas to the whole class.

To illustrate an impromptu oral task, teacher could elect a student and asked the student to summarize the day's lesson for example. Other students could help to 'fill-in' the gap the student may have missed out. In this way, others would also be given the opportunity to voice their opinions and ideas although they are not the one to give the summary.

9.3 Methods

9.3.1 Participants

Two secondary schools participated in the sub-study. One was identified as a high-performing school because it was randomly selected from the 50 best performing secondary schools that offered both Express and Normal Academic streams that was based on the year 1999-2002 GCE "O" Level Examination results released by the Ministry of Education (MOE), while the other secondary school was identified as a non-high performing school because it was not in this top 50 school ranking.

In each school, one Express and one Normal Academic stream classes in the secondary one level were chosen to participate in the study. This study adopted an experimental design method and therefore comparison classes were also chosen for the respective participating/experimental classes in each stream. The study needed to follow up the students for a period of 18 months, that is to say students were tracked as they moved from the Secondary One level to the Secondary Two level. The students from the non-high performing school were streamed at the end of their first year secondary academic year. As a result, some students in the experimental classes as well as in the respective comparison classes were lost due to streaming. In addition, the experimental teacher in the Express stream was replaced by another teacher at the beginning of the second year of the study because the former had taken long maternity leave. Table 9.3 presents a profile of the students and their respective mathematics teachers in this study with regard to gender.

Table 9.3. A profile of students and teachers

High-performing school	Express stream		Normal Academic stream	
	Experimental class	Comparison class	Experimental class	Comparison class
No. of Students	40	38	40	41
Male	18	19	24	24
Female	22	19	16	17
Math Teachers				
Gender	Female	Female	Female	Female
Length of teaching experience	2 yr 10 mth	8 yr	2 mth	8 yr
Qualification	PGDE	PGDE	PGDE	PGDE
Non-high-performing school				
No. of Students	38 (21)	39 (22)	39 (24)	39 (22)
Male	18 (10)	16 (9)	19 (10)	19 (12)
Female	20 (11)	23 (13)	20 (14)	20 (10)
Math teachers				
Gender	Female *	Female	Male	Male
Length of teaching experience	3 yr	10 yr	3 yr	30 yr
Qualification	PGDE	PGDE	PGDE	MEd

Note. The brackets indicate the number of students tracked in the second year of the study after students were streamed at the end of their secondary one. * In the second year, the teacher (who went on maternity leave) in the experimental class was replaced by another female teacher with 6 yrs of teaching experience and having the same qualification.

9.3.2 Instruments

The four instruments designed in this study were: questionnaire surveys, "journal writing task" tests, journal writing and oral presentation intervention worksheets, and interviews questions for teachers and students.

Questionnaire surveys

As reported in Chapter 2, there were pre- and post-questionnaire surveys, both intended to measure students' attitudes toward mathematics and learning of mathematics before and after 18 months of exposure to the use of communication tasks (see Appendices 2.2 and 2.3). Both the pre- and post-questionnaire surveys contained two parts; the first part consisted of 22 items which were constructed in the domain of attitudes toward mathematics to address factors that were considered important and they were in particular: (a) General view towards mathematics, (b) Anxiety towards learning of mathematics, (c) Feelings about their own mathematics performance, and (d) Belief towards the use of mathematics. In order to detect any possible changes in students' attitudes over a period of time, the same 22 items were used in the first part of the pre- and post questionnaire surveys

To respond to these items, students had to agree or disagree on a nine-point Likert-type scale format with the following anchors: *1 = Disagree totally, 2 = disagree a lot, 3 = Disagree, 4 = Disagree a little, 5 = Neither disagree nor agree, 6 = Agree a little, 7 = Agree, 8 = Agree a lot and 9 = Agree totally.*

The second part of the pre-questionnaire survey consisted of six items which were intended to find out students' pre-knowledge and experience regarding the 'new ways' of learning mathematics. For example, one item (which comprised of four sub items) was constructed to find out whether students had done a lot of writing and oral discussion in their previous mathematics learning. Similarly, students' responses were again based on six-point Likert-type scale (*1 = Almost every day, 2 = 2 or 3 times a week, 3 = Once a week or two weeks, 4 = Once a month, 5 = rarely, and 6 = never).*

On the contrary, the second part of the post-questionnaire survey consisted of 31 items (also see below) asking students from the experimental classes about their views toward communication tasks; more specifically, 15 items and 16 items were constructed to measure students' general feelings, belief, and perception about their own ability to engage in, journal writing task and oral presentation task respectively. Students' responses were based on the previously mentioned alike nine-point Likert-type scale format as found in the first part of the questionnaire survey.

To check the feasibility of the items in the pre-questionnaire survey, the initial version of the pre-questionnaire survey was piloted on 8/9 January 2004 on 56 secondary one students from two schools (who were not involved in the study). As a result, some of the items were rephrased and modified so that students could understand the questions better.

Both students from the experimental and comparison classes took the same pre-questionnaire survey. At the end of 18 months of intervention, students from the experimental classes took the post-questionnaire survey whereby the second part of the survey contained items measuring students' views toward communication tasks (as mentioned above), whereas the students in the comparison classes took the post-questionnaire survey with the second part of the survey containing items which were similar to those in the second part of the pre-questionnaire survey.

"Journal writing task" tests

There were pre- and post-"journal writing task" tests. Both tests contained two journal writing questions. The intention of the pre-"new strategy task" test was to examine and find out the extent to which students' writings illustrate the ability – to explain and use mathematical reasoning to show understanding of mathematical concepts, and to explain the use of strategies or procedures in solving problems. The questions in the post-"new strategy task" test were different but they were constructed to be parallel to those in the pre-test and the intention was to detect if there was any change in students' communication ability after 18 months of intervention.

A pilot study of the pre-"journal writing task" test was conducted on 19 February 2004 with 39 secondary one students; 19 students from the Normal Academic stream and 20 students from the Express stream of one school (not involved in this study). Generally, students had no difficulty understanding the questions and students took about 20 minutes to complete the test.

Journal writing and oral presentation intervention worksheets
One of the main objectives of the study was to find out how to effectively and efficiently integrate communication task into teachers' daily mathematics teaching instructions and students' learning. Thus, the study had systematically constructed appropriate journal writing and oral presentation prompts for teachers to implement in their own classrooms.

The worksheets were designed based on the teachers' scheme of work. At the beginning, the journal writing questions were constructed by the researcher however; subsequently teachers were encouraged to design their own questions. Due to the fact that journal writing was new to both the teachers and students, only journal writing tasks were used for the first nine months of the intervention period. In the second year of the study, the teachers then implemented oral presentation tasks for the next four months of the intervention period.

At times, the researcher was there to observe and video tape the teachers implementing the intervention tasks in their teaching. One of the aims was to provide the teachers feedback, advices and guidance about the use of the intervention tasks, as well as to communicate with the teachers and students constantly. The other reason was to gain more ideas from the classroom observations so that subsequent intervention tasks could be constructed or designed.

Teachers and students interviews
As said earlier, the interview questions (also see Appendices 2.4 and 2.5) were constructed with the intention of getting information from both teachers and students their experiences, understandings, comments, opinions and suggestions regarding the use of communication tasks as part of teaching instruction as well as learning process.

Interviews with teachers were done individually whereas interviews with students were done in groups of three. In each experimental class, six students (two high ability, two average ability and two low ability) were identified and chosen by the respective teachers. Therefore for each experimental class, two interview sessions were carried out with the students and in each session three students were interviewed.

9.3.3 Procedures and data collection

The pre-questionnaire survey was conducted on 12 February 2004 for all students in the experimental and comparison classes in the high performing school while the survey was conducted on the 13 February 2004 in the non-high performing school. 100% response rate was noted in the high performing school while 99.4% response rate was recorded in the non-high performing school. At the end of 18 months of intervention, the post-questionnaire survey was again conducted separately in each school. In the high performing school, the post-questionnaire survey was conducted after students' mid-year examination on the 20 May 2005 while in the non-high performing school it was administrated on 26 April 2005 (before students' mid-year examination). The response rate was noted to be 92.5% and 96.6% respectively in the high performing and non-high performing school. Both the questionnaire surveys were administrated and the answered scripts were collected by the respecting mathematics teacher in each class.

Likewise, the pre-"journal writing task" test was conducted to all students in the experimental and comparison classes on 9 and 10 March 2004 in the high performing school with a response rate of 99.4% and in non-high performing school with a 100% response rate respectively. The post-"journal writing task" test was conducted on the same day as the post-questionnaire survey mentioned above. The response rate was noted to be 90.6% and 96.6% respectively in the high and non-high performing school.

Both the pre- and post-tests were administrated and the answered scripts were collected by the respecting mathematics teacher in each class.

Part of the study also examined students' performance in their respective school mathematics examinations. One of the focuses was to investigate the impact of students' school mathematics results due to the exposure of 18 months of communication tasks in their daily mathematics learning. With the help of the participating mathematics teachers and their Heads of Department, the study collected students' Primary School Leaving Examination (PSLE) mathematics grades, School B004 mid-year mathematics examination scores (M2004), School B004 final-year mathematics examination scores (F2004), and School B005 mid-year mathematics examination scores (M2005).

During the 18 months of intervention, teachers integrated communication tasks in their daily mathematics classroom teaching. Every intervention task that students had taken was made-copies and filed accordingly. The Express stream students in both schools followed the same scheme of work (SOW), similarly, the Normal Academic stream students in both schools also worked on an identical SOW (but different from the Express stream). Although the participating teachers in both schools worked on a similar SOW in the same stream, different teachers have different styles and taught at different pace. Thus, the type of communication task prompts that students worked through and the total number of tasks worked on were different in different classes. On average, each experimental class had worked through about 20 communication tasks at the end of 18 months of intervention. At times, the researcher was there in the classroom to observe and take field-notes, or/and video record the lesson. The lesson observed might not necessarily be the lesson whereby a communication task was administrated. The researcher made regular classroom observations to for a variety of reasons. On average, researcher had made 13 classroom observations and four video recording in each classroom.

After all the post-questionnaire survey and post-"journal writing task" test were administrated, interviews with the participating teachers and participating students were arranged and conducted. The interviews with the teachers were conducted individually and each interview lasted for about 30-40 minutes and was audio recorded. The interview with the students were in groups of 3-4 students at a time and each interview lasted between 15-25 minutes and similarly, these interviews were audio recorded. All the interviews were done in the month of May 2005 except for one teacher in the non-high performing school, the interview was done near the end of April 2005.

9.3.4 Data process and analysis

The data collected from the pre- and post-questionnaire surveys were analyzed using descriptive statistics such as percentages was employed. In addition, Mann-Whitney U tests were also used to examine and detect if there was any significant differences in the responses to the questionnaire items between the experimental and comparison classes.

The pre- and post-"journal writing task" tests that were used to examine the effect of students' pre- and post communicative abilities comprised of two questions respectively. All the questions were graded using task-specific rubrics. The main focuses of the questions were looking into students' ability in "mathematical reasoning" and "presenting their writing coherently and systematically". Besides using descriptive statistic such as percentages to report on students' overall achievement in the abovementioned abilities before and after about one and a half year of using communication tasks, Mann-Whitney U tests were also employed to examine and detect if there were any significant differences in performance between the experimental and comparison classes. In addition, Wilcoxon Signed Ranks tests were also used to examine if there were any significant change in individual student's change in performance from the pre to the post-tests.

Students' PSLE math grades were used as an indicator to identify the experimental and comparison students' equivalent level in the beginning of the study and Chi-square test was used to test this level of equivalence. Subsequently students' three school academic math results (two results from the first year of intervention and one result from the second year of intervention) were collected and tracked. These three results were examined using T-test to see if there were any significant changes in students' school math achievement over time due to the exposure to communication tasks. Changes between the experimental and comparison classes, and changes within each class were both examined. More specifically, a 3×2 ANOVA with time (M2004, F2004, and M2005) as a within-subject factor and treatment (experimental vs. comparison) as a between subjects factor was used to detect the potential effect of the intervention program on the students from the experimental classes.

The interview data from the teachers and students were first recorded in audio form and then individually transcribed and documented. Using qualitative method, the data were categorized and described to find out about teachers' and students' views, feelings, attitudes and valuable suggestions and comments about using the communication tasks during learning and teaching of mathematics.

9.3.5 Limitations of the study

The study needed to follow- up with students and teachers for a period of about 18 months. During this long period of intervention, there were certain variables which were difficult to control. For example, in one school, the administration had decided to stream the Secondary One students at the end of the year which was also about the end of the 10^{th} month of the study. Thus, students in both the experimental and comparison classes were changed at the beginning of the second year of the study i.e. when students began their academic year of secondary two. As a result, students who had received or not received intervention right at the beginning of the study were all mixed up. Thus, students were 'lost' in both the experimental and comparison classes. However, the study managed to track about 50% of the original students (both in the experimental and comparison classes) from the beginning of the study.

In another case, the teacher in the experimental Express class was not able to follow-up with the class because she had to go on a long term maternity leave, and therefore the teacher had to be replaced. Therefore, the changes of administration policy in school regarding the issue of streaming and teacher transfer were some variables that were beyond the control for this research.

This study focused on the use of both journal writing and oral presentation tasks in mathematics learning. Ideally teachers should incorporate both the use of journal task and oral task at the beginning of the intervention period. However, due to the fact that both these tasks were new to teachers and students, teachers only implemented journal writing task into their daily teaching instruction during the first year of the study. It was only in the second year of the study that teachers then integrated oral task in their scheme of work. The second year of the intervention was a short period that lasted only about four to five months. In that sense, it would be helpful if the study could have been extended for further evaluation. The extension period would have given teachers and students a longer time to work on the oral tasks more systematically.

In addition, it would also have lead teachers and students work on both the tasks more consistently so that further revision and evaluation could be possible. Thus, time constraint was a limitation to the study.

Both journal writing and oral presentation tasks were considered as alternative moods of assessing students' understanding and learning of mathematics. However, during the period of intervention, none of the teachers' evaluation on students' work on these tasks was taken into account as part of students' final school achievement scores. Thus, the researchers believed that this could lead to some students not attempting the tasks seriously. Moreover, teachers were observed to make time from their daily, pre-planned schedule classroom teaching to do the new tasks and this often lead to teachers doing the new tasks only when they were reminded or told. The study hoped that teachers could have integrated the new tasks into their daily scheme of work more consistently and regularly.

9.4 Results and Discussions

As described above, the four main sets of data that were collected in this study, included students' responses to the pre- and post-questionnaire surveys which investigated students' attitudes toward mathematics and mathematics learning, school-based mathematics examination scores which traced students' academic performance in conventional mathematics tasks, pre- and post-"journal writing task" tests which were used to detect students' achievement in mathematical communication abilities, as well as interviews which intended to explore both teachers and students' perceptions about communication task.

In this section, the results are reported in sub-categories according to the types of data collected and in each sub-categories, it will report the results obtained first in the experimental Express class, the change within the class in the pre- and post-results, the change between the class and the comparison class, after that followed by the experimental Normal Academic class in the same sequence.

9.4.1 Questionnaire surveys

In this category, it would be further sub-divided into three sub-sections and in each of the sub-sections, it would first report the results with respect to the high performing school and then results from the non-high performing school. The three sub-sections were classified due to the different parts of the questionnaire surveys.

The first sub-section would report the results to the part A of the questionnaire surveys that talked about students' general attitudes toward mathematics and mathematics learning.

The second sub-section would report the results to the part B of the pre-questionnaire survey which all the students took and part B of the post-questionnaire survey that only the comparison classes took. The items in this part of the survey were intended to measure students' experiences with some of the new strategies in mathematics learning.

In the third sub-section, it would report the results to the part B of the post-questionnaire survey that only the experimental students took. The items in this part of the survey were constructed to measure experimental students' general attitudes toward the communication tasks.

Pre- and post-questionnaire surveys: Measuring students' general attitudes toward mathematics and mathematics learning
In the first part of the pre- and post-questionnaire surveys which was categorized as Part A consisted of 22 items. Six items were constructed to measure students' general view and liking about mathematics, students' anxiety towards learning of mathematics, students' perception about their own ability in doing mathematics respectively, and four items were constructed to measure students' general belief in the usefulness of mathematics. Specifically the item number and the description were as follows:

Q1G: I enjoy doing mathematics.
Q5G: Mathematics is hard for me.
Q9G: Mathematics is interesting to me.
Q13G: I don't have good feelings about mathematics.
Q16G: I like spending time on studying mathematics.
Q19G: I don't like to attend math lessons.
Q2A: I am never under a terrible strain in a math class.
Q6A: I am not afraid of doing mathematics.
Q10A: I am unable to think clearly when doing mathematics.
Q14A: I feel lost when trying to solve math problems.
Q17A: It makes me nervous to even think about having to do a math problem.
Q20A: I have a lot of confidence when it comes to mathematics.
Q3P: I am sure I can learn mathematics well.
Q7P: I can get good grades in mathematics.
Q11P: I am not good at mathematics.
Q15P: I don't think I can do well in mathematics.
Q21P: I like solving challenging math problems.
Q22P: I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself.
Q4B: I believe mathematics is useful.
Q8B: It is important to know mathematics nowadays.
Q12B: Studying mathematics is a waste of time.
Q18B: I will use mathematics a lot as an adult.

Note. Q12 reads as the question number 12 while the G, A, P and B represents the different aspects such as general view (G), anxiety (A), performance (P), and belief (B) respectively.

1. High-performing school

In the high-performing school (HPS), one Express stream class and one Normal Academic stream class were selected to take part in the study and these classes received intervention communication tasks and they were classified as experimental classes. Due to the design of the study, each experimental class also had a corresponding comparison class. Therefore in total, two experimental classes and two corresponding comparison classes took part in the questionnaire surveys. In this sub-section, the study would first report the results from the classes from the Express stream and then followed by the classes from the Normal Academic stream.

Express stream. To observe students' attitudes toward mathematics and mathematics learning before and after the intervention study, the percentages of responses in each item in the pre- and post-questionnaire surveys were evaluated and examined. In addition, in order to analyze if there was any differences in responses to the items in the questionnaire surveys between the experimental Express class and the comparison Express class, Mann-Whitney U tests were used to investigate this aspect.

Table 9.4 shows the percentages of the students in the experimental class responding to each of the question items and the items were classified accordingly to the domain of attitudes that addresses the four factors that were considered in the study to be important. The data revealed that by comparing the percentages of students responding to the pre-questionnaire and post-questionnaire survey, generally students had become less positive in their responses to almost all the items. For example, in item Q5G, by combining all the percentages starting from *Agree a little* to *Agree totally*, we observed that in the pre-questionnaire survey, 37.5% of the students had the same opinion that '*mathematics is hard for me'* whereas in the post-questionnaire survey, this percentage had increased to 57.5%, i.e. students seem to have felt mathematics was harder now compared to about 18 months ago. Similarly, by adopting the same way of adding the total percentages as mentioned above to represent the category of students as would generally 'agree' and total percentages from *Disagree to Disagree a little as* another category of students as would generally 'disagree', we would be able to make the following observation from the table.

To illustrate further, in item Q14A for example, there was a percentage increase from 30% (pre-test) to 55% (post-test) where students generally 'agree' to feel '*lost when trying to solve math problem'.* In the domain of students' perception of their own ability to do mathematics, item Q3P showed that there was an increase in percentage of students 'disagree' that '*I am sure I can learn mathematics well'* from 0% (pre-test) to 15% (post-test) and likewise, there was also an increase in percentages of students 'agree' that '*I am not good at mathematics'* from 30% (pre-questionnaire survey) to 45% (post-questionnaire survey). This phenomenon that students had generally became more negative in their general attitudes toward mathematics and learning of mathematics was not totally unexpected. As students moved to a higher level of learning, it was understandable that students would naturally feel that mathematics contents had became more difficult to grasp or learn and as a result, they would not like mathematics as before and feel less confident about their own learning.

Table 9.4. Distribution of responses for experimental express class (HPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1G	*pre*	0%	0%	5.0%	12.5%	15.0%	7.5%	25.0%	12.5%	22.5%
	post	0.0%	0.0%	2.5%	12.5%	20.0%	20.0%	20.0%	10.0%	15.0%
Q5G	*pre*	15.0%	0%	20%	15%	12.5%	27.5%	5%	5%	0%
	post	7.5%	2.5%	10.0%	12.5%	10.0%	25.0%	22.5%	7.5%	2.5%
Q9G	*pre*	0%	0%	5%	5%	12.5%	22.5%	12.5%	27.5%	15%
	post	0.0%	5.0%	2.5%	10.0%	17.5%	20.0%	17.5%	15.0%	12.5%
Q13G	*pre*	25%	2.5%	27.5%	10%	20%	7.5%	5%	2.5%	0%
	post	17.5%	15.0%	10.0%	7.5%	27.5%	17.5%	5.0%	0.0%	0.0%
Q16G	*pre*	0%	5%	2.5%	15%	27.5%	22.5%	10%	5%	12.5%
	post	0.0%	5.0%	10.0%	7.5%	22.5%	20.0%	25.0%	2.5%	7.5%
Q19G	*pre*	10%	22.5%	20%	12.5%	10%	10%	10%	2.5%	2.5%
	post	10.0%	15.0%	15.0%	12.5%	25.0%	15.0%	7.5%	0.0%	0.0%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2A	*pre*	2.5%	5.0%	17.5%	20.0%	17.5%	20%	10.0%	5.0%	2.5%
	post	5.0%	10.0%	22.5%	22.5%	10.0%	12.5%	10.0%	2.5%	5.0%
Q6A	*pre*	0%	0%	7.5%	12.5%	17.5%	10%	30%	30%	17.5%
	post	2.5%	0.0%	2.5%	10.0%	15.0%	17.5%	35.0%	7.5%	10.0%
Q10A	*pre*	7.5%	12.5%	25%	7.5%	20%	12.5%	10%	0%	5%
	post	5.0%	2.5%	27.5%	12.5%	20.0%	12.5%	15.0%	2.5%	2.5%
Q14A	*pre*	12.5%	10%	12.5%	7.5%	27.5%	20%	7.5%	0%	2.5%
	post	7.5%	2.5%	12.5%	5.0%	17.5%	32.5%	17.5%	2.5%	2.5%
Q17A	*pre*	12.5%	12.5%	17.5%	10%	12.5%	17.5%	10%	5%	2.5%
	post	7.5%	12.5%	30.0%	7.5%	20.0%	5.0%	0.0%	10.0%	7.5%
Q20A	*pre*	2.5%	7.5%	12.5%	10%	22.5%	7.5%	17.5%	15%	5%
	post	2.5%	7.5%	12.5%	15.0%	15.0%	20.0%	20.0%	5.0%	2.5%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3P	*pre*	0%	0%	0%	0%	27.5%	17.5%	25.0%	22.5%	7.5%
	post	2.5%	0.0%	5.0%	7.5%	22.5%	15.0%	27.5%	7.5%	12.5%

Q7P	*pre*	0%	0%	0%	5%	40%	12.5%	17.5%	17.5%	7.5%
	post	2.5%	2.5%	2.5%	10.0%	30.0%	22.5%	20.0%	2.5%	7.5%
Q11P	*pre*	17.5%	5%	12.5%	7.5%	27.5%	12.5%	12.5%	0%	5%
	post	5.0%	7.5%	7.5%	7.5%	27.5%	12.5%	20.0%	10.0%	2.5%
Q15P	*pre*	25%	0%	25%	15%	20%	12.5%	2.5%	0%	0%
	post	10.0%	5.0%	12.5%	20.0%	15.0%	15.0%	17.5%	2.5%	2.5%
Q21P	*pre*	0%	5%	15%	2.5%	15%	20%	7.5%	25%	10%
	post	5.0%	2.5%	20.0%	5.0%	17.5%	12.5%	15.0%	12.5%	10.0%
Q22P	*pre*	22.5%	7.5%	30%	5%	17.5%	7.5%	5%	0%	5%
	post	22.5%	10.0%	17.5%	22.5%	12.5%	7.5%	2.5%	5.0%	0.0%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4B	*pre*	0%	0%	0%	0%	5.0%	10.0%	32.5%	17.5%	35.0%
	post	0.0%	0.0%	0.0%	5.0%	5.0%	7.5%	37.5%	15.0%	30.0%
Q8B	*pre*	0%	0%	0%	0%	7.5%	2.5%	25%	22.5%	32.5%
	post	0.0%	0.0%	0.0%	2.5%	7.5%	7.5%	15.0%	27.5%	40.0%
Q12B	*pre*	37.5%	20%	20%	7.5%	10%	0%	2.5%	2.5%	0%
	post	25.0%	30.0%	20.0%	12.5%	7.5%	2.5%	2.5%	0.0%	0.0%
Q18B	*pre*	0%	0%	0%	0%	35%	5%	25%	17.5%	17.5%
	post	0.0%	2.5%	2.5%	0.0%	30.0%	7.5%	30.0%	17.5%	10.0%

Note. DT: Disagree totally; DAL: Disagree a lot; D: Disagree; DLL: Disagree a little; N: Neither disagree nor agree; ALL: Agree a little; A: Agree; AAL: Agree a lot; AT: Agree totally. (The same notations apply to the tables below)

Interestingly, in the domain of students' beliefs in the usefulness of mathematics and mathematics learning did not show much change in the students' responses to item Q12B and Q18B although there is a slight increase in students disagreeing that '*I believe mathematics is useful'* in Q4B (5% increment) and '*It is important to know mathematics nowadays'* Q8B (2.5% increment). The data also revealed that out of the 22 items, only five items reflected that students had become more positive in their attitudes after in the post-test compared to the results in the pre-test and these items are specifically item Q16G, Q19G, Q17A, Q20A and Q22P.

We would next examine the responses in the corresponding comparison Express class. The data in Table 9.5 revealed the students in the comparison Express class had responded more negatively to 15 items out of 22 items compared from the post to the pre-questionnaire survey based on the same calculation of total percentages explained above. Specifically, the items that showed improvement in attitudes were Q13G, Q19G, Q2A, Q10A, Q17A, Q4B and Q8B.

It is interesting to note that this class seemed to have less anxiety in their learning of mathematics compared to the experimental Express class whereby the experimental Express class had only one item in that domain that showed improvement in attitude after 18 months from the start of the study.

However, the changes in percentages from pre- to post-questionnaire survey in the above 7 (mentioned) items were not extensive enough bring about any significant results. For example in item Q13G, when 18.4% of students generally 'agree' that they '*don't have good feelings about mathematics'* in the pre-questionnaire survey, this percentage was dropped to about the same 18% in the post-test, only 0.4% change was detected.

As explained earlier, before the start of the study and after 18 months from the start of the study, students' attitudes toward mathematics and mathematics learning between the experimental Express class and the corresponding comparison Express class were both examined using pre- and post- questionnaire survey. The intention was to identify the question items that showed significant difference in students' responses between the classes. The study would like to examine whether exposure to communication tasks in the daily mathematics learning had any effect on students' attitudes toward mathematics and their mathematics learning compared to students' who were not exposed to such new tasks.

Table 9.5. Distribution of responses for comparison express class (HPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1G	*pre*	0%	0%	5.4%	5.4%	16.2%	5.4%	27%	13.5%	27%
	post	0.0%	0.0%	10.3%	7.7%	23.1%	10.3%	25.6%	7.7%	15.4%
Q5G	*pre*	10.5%	7.9%	21.1%	18.4%	18.4%	10.5%	10.5%	2.6%	0%
	post	0.0%	12.8%	12.8%	15.4%	28.2%	10.3%	20.5%	0.0%	0.0%
Q9G	*pre*	0%	0%	5.3%	2.6%	21.1%	13.2%	18.4%	10.5%	28.9%
	post	0.0%	2.6%	5.1%	2.6%	20.5%	23.1%	17.9%	15.4%	12.8%
Q13G	*pre*	23.7%	7.9%	18.4%	5.3%	26.3%	10.5%	5.3%	2.6%	0%
	post	17.9%	7.7%	20.5%	12.8%	23.1%	15.4%	2.6%	0.0%	0.0%
Q16G	*pre*	0%	2.6%	13.2%	5.3%	23.7%	15.8%	10.5%	10.5%	18.4%
	post	0.0%	0.0%	5.1%	7.7%	35.9%	17.9%	17.9%	2.6%	12.8%
Q19G	*pre*	13.2%	5.3%	28.9%	10.5%	34.2%	2.6%	5.3%	0%	0%
	Post	10.3%	10.3%	30.8%	5.1%	38.5%	5.1%	0.0%	0.0%	0.0%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2A	*pre*	2.6%	2.6%	2.6%	23.7%	21.1%	5.3%	13.2%	13.2%	13.2%
	post	5.1%	0.0%	10.3%	17.9%	20.5%	17.9%	15.4%	7.7%	5.1%
Q6A	*pre*	5.3%	0%	5.3%	10.5%	13.2%	13.2%	13.2%	15.8%	23.7%
	post	2.6%	0.0%	10.3%	10.3%	23.1%	12.8%	20.5%	10.3%	10.3%
Q10A	*pre*	13.2%	13.2%	7.9%	5.3%	31.6%	10.5%	10.5%	5.3%	2.6%
	post	2.6%	7.7%	20.5%	12.8%	33.3%	17.9%	5.1%	0.0%	0.0%
Q14A	*pre*	10.5%	15.8%	13.2%	10.5%	18.4%	18.4%	7.9%	5.3%	0%
	post	7.7%	15.4%	12.8%	12.8%	20.5%	23.1%	5.1%	0.0%	2.6%
Q17A	*pre*	18.4%	5.3%	15.8%	18.4%	18.4%	21.1%	2.6%	0%	0%
	post	10.3%	10.3%	15.4%	20.5%	30.8%	12.8%	0.0%	0.0%	0.0%
Q20A	*pre*	2.6%	5.3%	10.5%	10.5%	21.1%	13.2%	7.9%	10.5%	18.4%
	post	2.6%	0.0%	10.3%	15.4%	43.6%	7.7%	10.3%	7.7%	2.6%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3P	*pre*	2.6%	0%	5.3%	7.9%	7.9%	7.9%	26.3%	21.1%	18.4%
	post	2.6%	0.0%	0.0%	2.6%	30.8%	12.8%	35.9%	10.3%	5.1%
Q7P	*pre*	2.6%	2.6%	0%	5.3%	23.7%	10.5%	26.3%	10.5%	18.4%
	post	0.0%	0.0%	5.1%	17.9%	28.2%	12.8%	23.1%	7.7%	5.1%
Q11P	*pre*	15.8%	15.8%	10.5%	7.9%	21.1%	10.5%	10.5?	0%	7.9%
	post	5.1%	2.6%	7.7%	15.4%	30.8%	17.9%	17.9%	0.0%	2.6%
Q15P	*pre*	21.1%	7.9%	23.7%	15.8%	18.4%	5.3%	5.3%	2.6%	0%
	post	12.8%	7.7%	12.8%	17.9%	23.1%	20.5%	5.1%	0.0%	0.0%
Q21P	*pre*	0%	2.6%	7.9%	7.9%	13.2%	23.7%	7.9%	5.3%	31.6%
	post	0.0%	0.0%	7.7%	5.1%	38.5%	20.5%	10.3%	5.1%	12.8%
Q22P	*pre*	28.9%	7.9%	18.4%	7.9%	23.7%	10.5%	0%	2.6%	0%
	post	12.8%	7.7%	17.9%	12.8%	33.3%	10.3%	5.1%	0.0%	0.0%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4B	*pre*	0%	0%	0%	2.6%	7.9%	5.3%	34.2%	13.2%	36.8%
	post	0.0%	0.0%	2.6%	0.0%	2.6%	7.7%	38.5%	12.8%	35.9%
Q8B	*pre*	0%	0%	0%	0%	18.4%	10.5%	18.4%	7.9%	44.7%
	post	0.0%	0.0%	0.0%	0.0%	10.3%	7.7%	33.3%	23.1%	25.6%
Q12B	*pre*	42.1%	7.9%	28.9%	2.6%	15.8%	2.6%	0%	0%	0%
	post	25.6%	23.1%	25.6%	5.1%	15.4%	5.1%	0.0%	0.0%	0.0%

Q18B	*pre*	0%	5.3%	5.3%	0%	21.1%	13.2%	26.3%	10.5%	18.4%
	post	5.1%	0.0%	0.0%	2.6%	33.3%	7.7%	20.5%	15.4%	15.4%

At the beginning of the study, by adopting the Mann-Whitney test, the experimental and comparison classes did not show any significant difference in their responses to the questionnaire items. However in the post-questionnaire survey, significant differences in students' responses to item Q2A (U [40, 39] = 557.50, p = 0.027) and item Q14A (U [40, 39] = 553.00, p = 0.024) were found between the classes.

The changes reported above seemed to show that students in the comparison class were less anxiety towards learning of mathematics compared to the experimental class. According to the teacher in the experimental class, this group of students in the experimental class was unusually more serious and conscious about their learning compared to the comparison class, thus this may be the reason why the students in the experimental class were more anxious in their learning compared to hose from the comparison class. Or did the communication tasks make the students in the experimental class also more anxious in their learning since now they had to do more writing and oral in their mathematics lessons? This could possibly be a reason because the teacher did express that the emphasis of communication tasks had brought about the emphasis of the use of the English language and students had naturally became more careful about the way they explained their work in writing and speech, and thus this may had caused an anxiety in them. However, the study needed to consider and carefully analyze other data collected in order to make any conclusion to the above question.

Normal Academic stream. We will first examine the students' responses to the 22 items of Part A of the pre- and post-questionnaire surveys.

Tables 9.6 and 9.7 show the percentages of students' responses to each Likert-scale from the experimental Normal Academic class and the comparison Normal Academic class respectively. The data in Table 9.6 revealed that by comparing the total percentages (as mentioned before in the above when reporting for the experimental Express class) from the pre- and post-questionnaire surveys, the experimental class had generally become less positive in attitudes towards mathematics and learning of mathematics.

Specifically, there were four items that showed improvement in students' general views towards mathematics (Q1G, Q5G, Q9G, Q16G), three items (Q7P, Q11P, Q15P) reflected students' 'agreeing' to be more positive in their own ability to perform in mathematics, three items (Q8B, Q12B and Q18B) showed students' responses to have improved in their beliefs towards the usefulness of mathematics. To illustrate further for instance, we observed that 84.6% of students 'agreed' that they '*enjoy doing mathematics'* (Q1G) in the post-questionnaire survey compared to 77.5% in the pre-questionnaire survey, and there was an increase in percentages of students 'disagreeing' that '*I don't think I can do well in mathematics'* (Q15P) from 55% in the pre-questionnaire survey compared to 64% in the post-questionnaire survey.

Table 9.6. Distribution of responses for experimental normal academic class (HPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	0%	2.5%	5.0%	7.5%	7.5%	10.0%	27.5%	15.0%	25.0%
	post	5.1%	.0%	.0%	5.1%	5.1%	15.4%	33.3%	15.4%	20.5%
Q5	*pre*	7.5%	7.5%	17.5%	12.5%	15.0%	25.0%	10.0%	5.0%	0%
	post	12.8%	2.6%	10.3%	30.8%	12.8%	10.3%	15.4%	2.6%	2.6%
Q9	*pre*	2.5%	2.5%	2.5%	7.5%	15.0%	15.0%	22.5%	15.0%	17.5%
	post	7.7%	.0%	.0%	5.1%	12.8%	17.9%	15.4%	20.5%	20.5%
Q13	*pre*	22.5%	12.5%	17.5%	15.0%	12.5%	10.0%	5.0%	0%	5.0%
	post	17.9%	10.3%	23.1%	5.1%	28.2%	7.7%	2.6%	2.6%	2.6%
Q16.	*pre*	7.5%	5.0%	10.0%	10.0%	25.0%	7.5%	12.5%	10.0%	12.5%
	post	10.3%	.0%	7.7%	2.6%	30.8%	25.6%	12.8%	5.1%	5.1%
Q19	*pre*	23.1%	15.4%	15.4%	12.8%	12.8%	7.7%	5.0%	2.5%	5.0%
	post	25.6%	10.3%	12.8%	15.4%	20.5%	7.7%	2.6%	.0%	5.1%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	2.5%	7.5%	10.0%	12.5%	17.5%	7.5%	17.5%	7.5%	17.5%
	post	5.1%	12.8%	15.4%	7.7%	12.8%	25.6%	15.4%	.0%	5.1%
Q6	*pre*	0%	0%	12.5%	12.5%	10.0%	17.5%	17.5%	10.0%	20.0%
	post	7.7%	7.7%	7.7%	10.3%	15.4%	15.4%	15.4%	7.7%	12.8%
Q10	*pre*	5.0%	20.0%	10.0%	20.0%	10.0%	22.5%	0%	7.5%	5.0%
	post	7.7%	7.7%	5.1%	5.1%	28.2%	20.5%	20.5%	2.6%	2.6%
Q14	*pre*	15.0%	5.0%	15.0%	2.5%	12.5%	30.0%	2.5%	7.5%	10.0%
	post	10.5%	7.9%	7.9%	7.9%	23.7%	26.3%	2.6%	7.9%	5.3%
Q17	*pre*	22.5%	5.0%	10.0%	12.5%	15.0%	22.5%	2.5%	2.5%	7.5%
	post	15.4%	5.1%	12.8%	5.1%	23.1%	15.4%	12.8%	2.6%	7.7%
Q20	*pre*	2.6%	5.1%	10.3%	10.3%	20.5%	12.8%	20.5%	2.6%	15.4%
	post	2.6%	2.6%	5.3%	13.2%	26.3%	21.1%	18.4%	.0%	10.5%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	0%	0%	2.5%	2.5%	15.0%	17.5%	27.5%	7.5%	27.5%
	post	2.6%	.0%	.0%	7.7%	10.3%	12.8%	30.8%	23.1%	12.8%
Q7	*pre*	2.5%	0%	2.5%	10.0%	25.0%	22.5%	22.5%	7.5%	7.5%
	post	7.7%	.0%	.0%	5.1%	17.9%	23.1%	10.3%	17.9%	17.9%
Q11	*pre*	15.0%	10.0%	5.0%	10.0%	17.5%	10.0%	17.5%	5.0%	10.0%
	post	15.4%	2.6%	15.4%	12.8%	15.4%	10.3%	12.8%	5.1%	10.3%
Q15	*pre*	15.0%	7.5%	17.5%	15.0%	15.0%	12.5%	12.5%	0%	5.0%
	post	20.5%	7.7%	17.9%	17.9%	12.8%	12.8%	2.6%	2.6%	5.1%
Q21	*pre*	12.8%	0%	12.8%	10.3%	10.3%	12.8%	15.4%	7.7%	17.9%
	post	12.8%	5.1%	15.4%	12.8%	5.1%	15.4%	12.8%	5.1%	15.4%
Q22	*pre*	15.4%	17.9%	.6%	12.8%	28.2%	10.3%	5.1%	0%	7.7%
	post	12.8%	.0%	23.1%	7.7%	28.2%	15.4%	2.6%	2.6%	7.7%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	0%	0%	0%	0%	5.0%	10.0%	22.5%	15.0%	47.5%
	post	2.6%	.0%	.0%	2.6%	2.6%	15.4%	17.9%	23.1%	35.9%
Q8	*pre*	0%	0%	2.5%	5.0%	10.0%	10.0%	17.5%	15.0%	40.0%
	post	2.6%	.0%	.0%	.0%	5.1%	15.4%	20.5%	20.5%	35.9%
Q12	*pre*	47.5%	15.0%	7.5%	7.5%	15.0%	7.5%	0%	0%	0%
	post	41.0%	10.3%	30.8%	5.1%	10.3%	.0%	.0%	.0%	2.6%
Q18	*pre*	7.5%	2.5%	0%	10.0%	20.0%	7.5%	7.5%	10.0%	35.0%
	post	2.6%	.0%	5.1%	5.1%	25.6%	12.8%	17.9%	7.7%	23.1%

Correspondingly, the percentages of the comparison Normal Academic class students responding to the 22 items in the pre- and post-questionnaire surveys in Table 9.7

showed that there were 12 items out of 22 which students had made improvement in attitudes based on comparing the percentages from the pre to the post-questionnaire surveys.

The items were specifically, Q1G, Q5G that measured students' general views toward mathematics and mathematics learning, Q7P, Q11P that measured students' beliefs in their own performance in mathematics, Q4B, Q8B, Q18B that measured students' beliefs in the usefulness in mathematics, and Q2A, Q6A, Q10A, Q14A and Q17A that measured students' anxiety towards learning of mathematics.

Interestingly, the data revealed that students in this class seemed to have become less anxious in their learning of mathematics when they had progressed to secondary two from secondary one, 18 months after the start of the study. What was captivating was that the teacher mentioned that the students of this class was always not very interested in their study be it mathematics or any other subject, and therefore the teacher was not very surprised to learn that her students were not anxious about their learning of mathematics. In the similar way, before the start of the study and after 18 months from the start from the study, comparisons of students' attitudes toward mathematics and mathematics learning between the experimental Normal Academic class and the comparison class were both examined.

As we can see from Table 9.7, at the beginning of the study, the experimental and the comparison classes did not show any significant difference in their responses to the 22 items mentioned above. However, in the post-questionnaire survey, significant difference in students' responses to item Q15P (U [39, 30] = 406.00, p = 0.028) was detected; there was significantly more percentages of students in the experimental class (64%) believed that they '*can do well in mathematics'* than those in the comparison class (36.6%). Although the teacher did mention that students in the experimental class was not always very serious about their learning, she believed that this experimental class students were more capable and they also seemed to have more confidence in themselves, compared to the comparison class.

Table 9.7. Distribution of responses for comparison normal academic class (HPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	9.80%	2.40%	4.90%	14.60%	2.40%	19.50%	22.00%	9.80%	14.60%
	post	6.70%	3.30%	6.70%	10.00%	3.30%	23.30%	20.00%	6.70%	20.00%
Q5	*pre*	0%	2.40%	17.10%	2.40%	22.00%	43.90%	12.20%	0%	0%
	post	13.30%	0.00%	6.70%	13.30%	13.30%	30.00%	13.30%	3.30%	6.70%
Q9	*pre*	2.40%	0%	4.90%	14.60%	7.30%	34.10%	17.10%	9.80%	9.80%
	post	3.30%	0.00%	3.30%	6.70%	23.30%	20.00%	26.70%	3.30%	13.30%
Q13	*pre*	9.80%	9.80%	19.50%	14.60%	22.00%	19.50%	2.40%	2.40%	0%
	post	13.30%	16.70%	13.30%	3.30%	23.30%	20.00%	6.70%	0.00%	3.30%
Q16.	*pre*	4.90%	7.30%	12.20%	14.60%	14.60%	24.40%	4.90%	9.80%	7.30%
	post	6.70%	6.70%	0.00%	3.30%	36.70%	20.00%	13.30%	10.00%	3.30%
Q19	*pre*	17.10%	12.20%	12.20%	22.00%	17.10%	7.30%	7.30%	2.40%	2.40%
	post	10.30%	10.30%	20.70%	3.40%	27.60%	17.20%	3.40%	0.00%	6.90%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	4.90%	2.40%	19.50%	19.50%	12.20%	19.50%	19.50%	0%	2.40%
	post	6.70%	0.00%	20.00%	10.00%	10.00%	33.30%	10.00%	3.30%	6.70%
Q6	*pre*	0%	0%	9.80%	19.50%	22.00%	17.10%	19.50%	7.30%	4.90%
	post	3.30%	3.30%	6.70%	10.00%	20.00%	13.30%	26.70%	10.00%	6.70%
Q10	*pre*	2.40%	2.40%	19.50%	14.60%	19.50%	24.40%	12.20%	2.40%	2.40%
	post	10.30%	3.40%	6.90%	20.70%	17.20%	17.20%	20.70%	3.40%	0.00%
Q14	*pre*	4.90%	0%	14.60%	7.30%	22.00%	41.50%	9.80%	0%	0%

	post	3.30%	3.30%	16.70%	10.00%	13.30%	30.00%	13.30%	3.30%	6.70%
Q17	*pre*	0%	9.80%	17.10%	14.60%	26.80%	22.00%	7.30%	0%	2.40%
	post	3.30%	10.00%	23.30%	10.00%	20.00%	23.30%	3.30%	0.00%	6.70%
Q20	*pre*	0%	2.40%	4.90%	14.60%	26.80%	31.70%	12.20%	7.30%	0%
	post	10.30%	0.00%	10.30%	6.90%	27.60%	20.70%	17.20%	0.00%	6.90%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	0%	2.40%	0%	9.80%	4.90%	24.40%	39.00%	9.80%	9.80%
	post	3.30%	0.00%	10.00%	6.70%	0.00%	30.00%	36.70%	3.30%	10.00%
Q7	*pre*	2.40%	0%	2.40%	9.80%	29.30%	29.30%	22.00%	0%	2.40%
	post	10.00%	3.30%	0.00%	3.30%	20.00%	26.70%	20.00%	13.30%	3.30%
Q11	*pre*	2.40%	4.90%	14.60%	14.60%	29.30%	12.20%	17.10%	2.40%	2.40%
	post	6.90%	3.40%	13.80%	13.80%	17.20%	17.20%	10.30%	10.30%	6.90%
Q15	*pre*	0%	17.10%	19.50%	9.80%	31.70%	12.20%	7.30%	2.40%	0%
	post	3.30%	3.30%	20.00%	10.00%	26.70%	16.70%	13.30%	3.30%	3.30%
Q21	*pre*	12.20%	0%	9.80%	12.20%	14.60%	22.00%	17.10%	7.30%	4.90%
	post	10.70%	0.00%	10.70%	17.90%	17.90%	10.70%	25.00%	3.60%	3.60%
Q22	*pre*	9.80%	2.40%	22.00%	14.60%	22.00%	7.30%	7.30%	2.40%	12.20%
	post	10.30%	6.90%	6.90%	6.90%	34.50%	17.20%	6.90%	0.00%	10.30%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	0%	0%	0%	0%	7.30%	4.90%	46.30%	17.10%	24.40%
	post	0.00%	0.00%	0.00%	0.00%	6.70%	13.30%	26.70%	23.30%	30.00%
Q8	*pre*	0%	0%	0%	2.40%	4.90%	2.40%	39.00%	19.50%	31.70%
	post	0.00%	0.00%	0.00%	0.00%	3.30%	10.00%	30.00%	30.00%	26.70%
Q12	*pre*	34.10%	14.60%	14.60%	14.60%	19.50%	2.40%	0%	0%	0%
	post	26.70%	30.00%	13.30%	0.00%	26.70%	0.00%	0.00%	0.00%	3.30%
Q18	*pre*	2.40%	0%	0%	7.30%	31.70%	4.90%	31.70%	17.10%	4.90%
	post	3.30%	0.00%	0.00%	6.70%	20.00%	13.30%	26.70%	16.70%	13.30%

2. Non-high-performing school

Likewise in the non-high-performing school (NHPS), one class from the Express stream and one class from the Normal Academic stream were selected to participate in this study and because these classes received intervention communication tasks, they were classified as experimental classes. Correspondingly, each experimental class was attached with one comparison class. In total, there were two experimental classes with two corresponding comparison classes which took part in the questionnaire surveys. This sub-section will first report the results obtained from the Express stream classes followed by the results obtained from the Normal Academic stream classes.

Express stream. In a similar sense, by observing Table 9.8 and considering the total percentages of students ‘agreeing’ and ‘disagreeing’, the data revealed that out of the 22 items, the experimental Express class had responded more positively to 13 items in the post-questionnaire survey compared to the pre-questionnaire survey. They were specifically Q5G, Q9G, Q13G, Q16G, Q6A, Q10A, Q17A, Q3P, Q15P, Q21P, Q22P, Q4B and Q12B.

It was obvious that in these two categories: *general views about mathematics* and students’ beliefs in their own *ability to perform mathematics,* received more items in which students’ had more positive responses. For examples in item Q5G “*mathematics is hard for me”*, the percentages of students ‘agreeing’ was dropped from 47.4% to 28.6%; in item Q9G “*mathematics is interesting to me”,* the percentages of students ‘disagreeing’ was dropped from 21.1% to 9.5%; in item Q13G “*I don’t have good feelings about mathematics”* had an increase in percentages of students ‘disagreeing’ from 57.9% to

61.8%; and in item Q16G "*I like spending time on studying mathematics"* saw an increase in percentages of students 'agreeing' from 34.2% to 38.1%.

Similarly, for the items that measure student' own beliefs in their ability to *perform mathematics* for example in item Q3P "*I am sure I can learn mathematics well"* the percentages of students 'agreeing' to increased from 60.5% to 66.7%; in item Q15P "*I don't think I can do well in mathematics"* also saw an increased in students 'disagreeing' from 44.8% to 52.5%; in item Q21P "*I like solving challenging math problems"* had an increase in percentages of students 'agreeing' from 34.2% to 57.1%; and for item Q22P "*I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself"* also had an increase in percentages of students 'disagreeing' from 44.8% to 66.6%.

Besides the abovementioned two categories, students also responded positively to three items out of six measuring students' *anxiety towards mathematics learning*, and two out of four items measuring students' *beliefs in the usefulness of mathematics.*

Table 9.8. Distribution of responses for experimental express class (NHPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	5.3%	2.6%	0.0%	2.6%	21.1%	10.5%	23.7%	18.4%	15.8%
	post	5.0%	0.0%	0.0%	10.0%	25.0%	10.0%	25.0%	15.0%	10.0%
Q5	*pre*	10.5%	7.9%	2.6%	2.6%	28.9%	15.8%	15.8%	10.5%	5.3%
	post	4.8%	0.0%	9.5%	23.8%	33.3%	9.5%	4.8%	4.8%	9.5%
Q9	*pre*	7.9%	0.0%	7.9%	5.3%	18.4%	13.2%	23.7%	10.5%	13.2%
	post	0.0%	0.0%	0.0%	9.5%	33.3%	9.5%	23.8%	9.5%	14.3%
Q13	*pre*	26.3%	7.9%	18.4%	5.3%	18.4%	10.5%	2.6%	0.0%	10.5%
	post	19.0%	23.8%	19.0%	0.0%	23.8%	9.5%	4.8%	0.0%	0.0%
Q16.	*pre*	13.2%	0.0%	10.5%	7.9%	34.2%	15.8%	10.5%	0.0%	7.9%
	post	9.5%	0.0%	4.8%	4.8%	42.9%	23.8%	9.5%	4.8%	0.0%
Q19	*pre*	31.6%	5.3%	18.4%	13.2%	13.2%	10.5%	0.0%	5.3%	2.6%
	post	19.0%	19.0%	9.5%	0.0%	42.9%	9.5%	0.0%	0.0%	0.0%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	10.5%	0.0%	2.6%	28.9%	18.4%	15.8%	10.5%	5.3%	7.9%
	post	9.5%	4.8%	9.5%	19.0%	14.3%	19.0%	9.5%	9.5%	4.8%
Q6	*pre*	10.5%	0.0%	10.5%	15.8%	18.4%	18.4%	13.2%	7.9%	5.3%
	post	9.5%	0.0%	0.0%	9.5%	19.0%	23.8%	23.8%	4.8%	9.5%
Q10	*pre*	13.2%	5.3%	10.5%	10.5%	21.1%	26.3%	5.3%	2.6%	5.3%
	post	0.0%	19.0%	14.3%	19.0%	19.0%	14.3%	9.5%	4.8%	0.0%
Q14	*pre*	15.8%	0.0%	13.2%	13.2%	15.8%	26.3%	7.9%	0.0%	7.9%
	post	0.0%	19.0%	14.3%	0.0%	23.8%	28.6%	4.8%	0.0%	9.5%
Q17	*pre*	13.2%	0.0%	7.9%	5.3%	34.2%	15.8%	10.5%	2.6%	10.5%
	post	4.8%	14.3%	14.3%	4.8%	23.8%	28.6%	0.0%	0.0%	9.5%
Q20	*pre*	7.9%	2.6%	10.5%	7.9%	31.6%	13.2%	13.2%	5.3%	7.9%
	post	14.3%	4.8%	4.8%	14.3%	23.8%	23.8%	9.5%	4.8%	0.0%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	5.3%	0.0%	2.6%	2.6%	28.9%	18.4%	21.1%	10.5%	10.5%
	post	0.0%	0.0%	0.0%	0.0%	33.3%	14.3%	23.8%	23.8%	4.8%
Q7	*pre*	7.9%	0.0%	7.9%	2.6%	39.5%	10.5%	15.8%	7.9%	7.9%
	post	9.5%	0.0%	0.0%	9.5%	42.9%	14.3%	19.0%	4.8%	0.0%
Q11	*pre*	13.2%	5.3%	5.3%	0.0%	28.9%	15.8%	13.2%	5.3%	13.2%
	post	0.0%	0.0%	14.3%	23.8%	14.3%	23.8%	4.8%	4.8%	14.3%
Q15	*pre*	13.2%	2.6%	13.2%	15.8%	28.9%	13.2%	5.3%	0.0%	7.9%
	post	4.8%	4.8%	38.1%	4.8%	28.6%	9.5%	0.0%	0.0%	9.5%
Q21	*pre*	10.5%	15.8%	2.6%	15.8%	21.1%	13.2%	2.6%	2.6%	15.8%

	post	9.5%	0.0%	4.8%	4.8%	23.8%	33.3%	14.3%	0.0%	9.5%
Q22	*pre*	13.2%	7.9%	10.5%	13.2%	23.7%	13.2%	2.6%	2.6%	13.2%
	post	19.0%	14.3%	23.8%	9.5%	14.3%	4.8%	4.8%	0.0%	9.5%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	2.6%	0.0%	0.0%	0.0%	18.4%	7.9%	26.3%	13.2%	31.6%
	post	0.0%	0.0%	0.0%	0.0%	23.8%	4.8%	23.8%	23.8%	23.8%
Q8	*pre*	2.6%	0.0%	0.0%	0.0%	15.8%	10.5%	26.3%	7.9%	36.8%
	post	0.0%	0.0%	4.8%	0.0%	19.0%	9.5%	23.8%	28.6%	14.3%
Q12	*pre*	44.7%	10.5%	13.2%	7.9%	18.4%	0.0%	0.0%	0.0%	5.3%
	post	57.1%	4.8%	14.3%	0.0%	19.0%	4.8%	0.0%	0.0%	0.0%
Q18	*pre*	2.6%	0.0%	0.0%	0.0%	39.5%	15.8%	21.1%	5.3%	15.8%
	post	0.0%	0.0%	0.0%	0.0%	52.4%	19.0%	14.3%	4.8%	9.5%

It was not surprising that students in this class had seemed generally ‘improved’ in their attitudes toward mathematics and mathematics learning because the changed of the teacher starting in the second year of intervention was believed to have made a positive impact on the students. During classroom observations, it was noted that the teacher was more encouraging and her expectation were more explicitly made known to students. Comparing to the teacher who first started out the intervention with the class, the teacher who took over in the second year was also noted to be more experience in her in teaching and classroom management.

Next, let us examine the responses from the corresponding comparison Express class.

Table 9.9 revealed that by comparing the pre- and post-results, this class had generally decreased in percentages of students responding positively to the questionnaire items. Only two items Q22P the measured students’ own belief in their *‘ability to perform mathematics’* saw an increase in the percentages of students ‘disagreeing’ to this negative item from 55.3% to 63.6%, and item Q18B that measured students’ ‘*beliefs in the usefulness in mathematics’* also saw an increase in percentages of students ‘agreeing’ to this positive item from 55.4% to 59.1%. All the 20 items beside the two items mentioned above received more negative responses in the post-questionnaire survey compared to the pre-questionnaire survey.

In addition, this study also looked at the initial stage of the experimental students’ attitudes toward mathematics and mathematics learning before the start of the 18 months of intervention and after the 18 months of exposure to communication tasks in the mathematics classroom. First by examining the pre-questionnaire survey between the experimental and the corresponding comparison classes, the Mann-Whitney tests revealed that there was a significant difference in students’ responses to question item Q6A “*I am not afraid of doing mathematics”* (U [38, 38] = 499.00, $p < .05$) which found that the comparison class was less ‘*afraid* ‘in doing mathematics compared to the experimental class.

This phenomenon was quite interesting because although (as mentioned earlier) the teacher in the experimental class was more encouraging and assuring but she was also very strict and firm with her students about work to be done properly. She certainly did not accept any sloppy work from her students and was very stern about classroom discipline.

In the post-questionnaire survey, no significant difference in the students’ responses between the experimental and comparison classes was found in all the items.

Table 9.9. Distribution of responses for comparison express class (NHPS)

General Views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	5.3%	0.0%	0.0%	2.6%	21.1%	15.8%	15.8%	21.1%	18.4%
	post	4.5%	4.5%	4.5%	13.6%	18.2%	9.1%	9.1%	13.6%	22.7%
Q5	*pre*	15.8%	5.3%	15.8%	7.9%	15.8%	15.8%	21.1%	0.0%	2.6%
	post	4.5%	0.0%	4.5%	18.2%	22.7%	18.2%	9.1%	13.6%	9.1%
Q9	*pre*	7.9%	0.0%	0.0%	7.9%	15.8%	13.2%	10.5%	26.3%	18.4%
	post	9.1%	4.5%	9.1%	4.5%	9.1%	13.6%	9.1%	13.6%	27.3%
Q13	*pre*	23.7%	10.5%	13.2%	2.6%	13.2%	28.9%	2.6%	2.6%	2.6%
	post	18.2%	13.6%	18.2%	9.1%	4.5%	9.1%	9.1%	4.5%	13.6%
Q16.	*pre*	2.6%	2.6%	15.8%	18.4%	13.2%	13.2%	18.4%	7.9%	7.9%
	post	9.1%	9.1%	13.6%	13.6%	13.6%	18.2%	4.5%	13.6%	4.5%
Q19	*pre*	13.2%	15.8%	15.8%	5.3%	23.7%	13.2%	7.9%	0.0%	5.3%
	post	18.2%	13.6%	9.1%	9.1%	13.6%	9.1%	9.1%	4.5%	13.6%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	7.9%	0.0%	15.8%	15.8%	15.8%	13.2%	7.9%	15.8%	7.9%
	post	13.6%	9.1%	22.7%	4.5%	13.6%	9.1%	18.2%	0.0%	9.1%
Q6	*pre*	0.0%	0.0%	5.3%	15.8%	21.1%	10.5%	7.9%	18.4%	21.1%
	post	9.1%	4.5%	4.5%	9.1%	22.7%	9.1%	22.7%	13.6%	4.5%
Q10	*pre*	10.5%	10.5%	5.3%	21.1%	31.6%	10.5%	2.6%	5.3%	2.6%
	post	4.5%	13.6%	9.1%	27.3%	13.6%	22.7%	4.5%	0.0%	4.5%
Q14	*pre*	7.9%	5.3%	10.5%	7.9%	15.8%	31.6%	7.9%	0.0%	13.2%
	post	4.5%	4.5%	13.6%	0.0%	27.3%	31.8%	4.5%	4.5%	9.1%
Q17	*pre*	10.5%	10.5%	15.8%	0.0%	26.3%	13.2%	7.9%	5.3%	10.5%
	post	0.0%	4.5%	22.7%	18.2%	22.7%	18.2%	4.5%	4.5%	4.5%
Q20	*pre*	5.3%	0.0%	5.3%	18.4%	23.7%	18.4%	18.4%	7.9%	2.6%
	post	0.0%	4.5%	18.2%	27.3%	4.5%	22.7%	9.1%	9.1%	4.5%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	0.0%	0.0%	0.0%	2.6%	13.2%	26.3%	28.9%	10.5%	18.4%
	post	0.0%	0.0%	0.0%	4.5%	22.7%	18.2%	27.3%	9.1%	18.2%
Q7	*pre*	0.0%	0.0%	2.6%	23.7%	15.8%	10.5%	34.2%	13.2%	0.0%
	post	4.5%	0.0%	18.2%	13.6%	18.2%	9.1%	22.7%	9.1%	4.5%
Q11	*pre*	7.9%	5.3%	13.2%	10.5%	18.4%	13.2%	23.7%	2.6%	5.3%
	post	4.5%	9.1%	18.2%	0.0%	18.2%	22.7%	9.1%	9.1%	9.1%
Q15	*pre*	10.5%	7.9%	18.4%	15.8%	21.1%	15.8%	7.9%	0.0%	2.6%
	post	9.1%	9.1%	22.7%	18.2%	4.5%	13.6%	9.1%	4.5%	9.1%
Q21	*pre*	7.9%	0.0%	10.5%	13.2%	21.1%	10.5%	5.3%	13.2%	18.4%
	post	9.1%	9.1%	18.2%	9.1%	4.5%	13.6%	9.1%	27.3%	0.0%
Q22	*pre*	10.5%	21.1%	7.9%	15.8%	28.9%	5.3%	5.3%	0.0%	5.3%
	post	18.2%	22.7%	22.7%	0.0%	18.2%	9.1%	4.5%	0.0%	4.5%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	0.0%	0.0%	0.0%	0.0%	5.3%	7.9%	21.1%	21.1%	44.7%
	post	0.0%	0.0%	0.0%	0.0%	9.1%	13.6%	31.8%	18.2%	27.3%
Q8	*pre*	0.0%	0.0%	0.0%	0.0%	7.9%	7.9%	18.4%	28.9%	36.8%
	post	0.0%	0.0%	0.0%	0.0%	9.1%	18.2%	13.6%	27.3%	31.8%
Q12	*pre*	31.6%	21.1%	18.4%	2.6%	23.7%	0.0%	2.6%	0.0%	0.0%
	post	36.4%	18.2%	13.6%	9.1%	9.1%	9.1%	0.0%	4.5%	0.0%
Q18	*pre*	5.3%	0.0%	0.0%	5.3%	34.2%	13.2%	21.1%	13.2%	7.9%
	post	4.5%	4.5%	0.0%	0.0%	31.8%	13.6%	18.2%	18.2%	9.1%

Normal Academic stream. Table 9.10 showed the percentages of students responding in the different level of the likert-scales of each of the question items. The data revealed that the experimental Normal Academic class had responded more positively to all the four items measuring students' '*beliefs in the usefulness of mathematics'* in the post-questionnaire survey than the pre-questionnaire survey; more positive responses were also found in the categories measuring students' '*general views towards mathematics'* specifically five items out of six: Q1G, Q5G, Q9G, Q13G and Q19G, and in the category measuring students' own '*beliefs in their ability to perform mathematics'* with four items: Q3P, Q7P, Q11P and Q15P; but only two items out of six items in the '*anxiety'* category. To illustrate, examine the category under *'performance'* for examples, 78.3% of students believed that they could "*learn mathematics well"* (Q3P) in the pre-results was increased to 87.5%; in item Q7P "*I can get good grades in mathematics"* also saw an increase in percentages of students 'agreeing' from 43.6% to 62.5%; in the negative item Q11P "*I am not good at mathematics"* had an increase in students 'disagreeing' from 28.2% to 33.4%; and for negative item Q15P also saw an increase in percentages of students 'disagreeing' from 46.1% to 54.1%. Thus, 11 items out of 22 received more positive responses from students in the post-questionnaire survey compared to the pre-questionnaire survey.

Once again, it was quite expected that students in this class were generally quite positive in their learning. The reason may be due to the teacher; as the teacher in this experimental class had excellent rapport with his students. He was observed to be very supportive and give confidence to his students and he always encouraged his students to have self-positive attitudes towards their own learning. It was not unusual to hear the teacher said: "You must believe that you can do it because it is not difficult!" even when students were first engaged in journal writing tasks.

The corresponding comparison Normal Academic class had nine items out of 22 items that showed increased in percentages of students responding more positively in the post than the pre-questionnaire surveys (see Table 9.11). They were specifically items Q16G, Q19G, Q6A, Q20A, Q7P, Q15P, Q22P, Q8B and Q18B. For examples in item Q7P "*I can get good grades in mathematics"* saw an increase in percentages of students 'agreeing' from 41% to 60%; consistently in negative item Q15P "*I don't think I can do well in mathematics"* there is also an increase in percentages of students 'disagreeing' from 43.7% to 63.2%; and for negative item Q22P also saw an increase in students 'disagreeing' from 33.4% to 45%. However, these three items were the only three items out of six items in the category that measured students' '*general views toward mathematics*' responded positively. In fact, students had responded more positively in the post-questionnaire compared to the pre-questionnaire in only 50% or less number of items in each of the four categories.

Table 9.10. Distribution of responses for experimental normal academic class (NHPS)

General views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	5.1%	0.0%	7.7%	15.4%	5.1%	2.6%	35.9%	12.8%	15.4%
	post	8.3%	0.0%	0.0%	4.2%	4.2%	20.8%	25.0%	12.5%	25.0%
Q5	*pre*	5.1%	5.1%	10.3%	12.8%	17.9%	15.4%	23.1%	2.6%	7.7%
	post	12.5%	4.2%	12.5%	8.3%	29.2%	8.3%	16.7%	4.2%	4.2%
Q9	*pre*	2.6%	2.6%	2.6%	15.4%	5.1%	12.8%	25.6%	15.4%	17.9%
	post	0.0%	0.0%	0.0%	4.2%	12.5%	20.8%	20.8%	12.5%	29.2%
Q13	*pre*	20.5%	12.8%	17.9%	17.9%	12.8%	7.7%	5.1%	5.1%	0.0%
	post	37.5%	4.2%	20.8%	8.3%	8.3%	12.5%	0.0%	4.2%	4.2%
Q16.	*pre*	2.6%	2.6%	10.3%	12.8%	15.4%	17.9%	23.1%	12.8%	2.6%
	post	4.2%	4.2%	12.5%	8.3%	37.5%	4.2%	12.5%	12.5%	4.2%
Q19	*pre*	17.9%	15.4%	23.1%	15.4%	17.9%	7.7%	0.0%	2.6%	0.0%
	post	25.0%	4.2%	37.5%	12.5%	12.5%	4.2%	0.0%	0.0%	4.2%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	17.9%	2.6%	10.3%	15.4%	10.3%	15.4%	17.9%	10.3%	0.0%
	post	16.7%	0.0%	4.2%	25.0%	12.5%	29.2%	4.2%	0.0%	8.3%
Q6	*pre*	2.6%	0.0%	2.6%	15.4%	30.8%	12.8%	20.5%	7.7%	7.7%

		DT	DAL	D	DLL	N	ALL	A	AAL	AT
	post	8.3%	0.0%	8.3%	8.3%	12.5%	8.3%	16.7%	16.7%	20.8%
Q10	*pre*	10.3%	2.6%	15.4%	12.8%	17.9%	28.2%	7.7%	2.6%	2.6%
	post	12.5%	0.0%	12.5%	4.2%	25.0%	12.5%	12.5%	4.2%	16.7%
Q14	*pre*	7.7%	12.8%	7.7%	2.6%	12.8%	17.9%	23.1%	12.8%	2.6%
	post	12.5%	4.2%	20.8%	4.2%	16.7%	29.2%	4.2%	0.0%	8.3%
Q17	*pre*	2.6%	7.7%	20.5%	7.7%	28.2%	20.5%	7.7%	2.6%	2.6%
	post	8.3%	8.3%	29.2%	0.0%	12.5%	25.0%	8.3%	0.0%	8.3%
Q20	*pre*	5.1%	5.1%	5.1%	23.1%	7.7%	15.4%	17.9%	15.4%	5.1%
	post	12.5%	4.2%	0.0%	8.3%	25.0%	16.7%	8.3%	16.7%	8.3%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	0.0%	0.0%	2.6%	7.7%	10.3%	15.4%	23.1%	17.9%	23.1%
	post	4.2%	0.0%	0.0%	4.2%	4.2%	12.5%	33.3%	16.7%	25.0%
Q7	*pre*	7.7%	2.6%	7.7%	12.8%	25.6%	20.5%	10.3%	12.8%	0.0%
	post	8.3%	0.0%	8.3%	0.0%	20.8%	12.5%	25.0%	8.3%	16.7%
Q11	*pre*	5.1%	7.7%	12.8%	2.6%	17.9%	15.4%	20.5%	5.1%	12.8%
	post	12.5%	4.2%	12.5%	4.2%	25.0%	20.8%	8.3%	0.0%	12.5%
Q15	*pre*	7.7%	5.1%	20.5%	12.8%	15.4%	17.9%	12.8%	2.6%	5.1%
	post	16.7%	8.3%	20.8%	8.3%	20.8%	12.5%	4.2%	4.2%	4.2%
Q21	*pre*	7.7%	7.7%	10.3%	12.8%	20.5%	15.4%	15.4%	5.1%	5.1%
	post	12.5%	4.2%	8.3%	16.7%	20.8%	4.2%	12.5%	8.3%	12.5%
Q22	*pre*	12.8%	7.7%	17.9%	10.3%	23.1%	5.1%	10.3%	0.0%	12.8%
	post	8.3%	8.3%	20.8%	16.7%	4.2%	8.3%	25.0%	4.2%	4.2%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	0.0%	0.0%	0.0%	0.0%	2.6%	5.1%	23.1%	23.1%	46.2%
	post	0.0%	0.0%	0.0%	0.0%	4.2%	0.0%	20.8%	20.8%	54.2%
Q8	*pre*	0.0%	0.0%	2.6%	0.0%	2.6%	7.7%	28.2%	10.3%	48.7%
	post	0.0%	0.0%	0.0%	0.0%	0.0%	4.2%	29.2%	20.8%	45.8%
Q12	*pre*	48.7%	12.8%	28.2%	2.6%	5.1%	2.6%	0.0%	0.0%	0.0%
	post	62.5%	16.7%	12.5%	0.0%	8.3%	0.0%	0.0%	0.0%	0.0%
Q18	*pre*	5.1%	0.0%	2.6%	2.6%	20.5%	10.3%	28.2%	7.7%	23.1%
	post	0.0%	0.0%	0.0%	0.0%	41.7%	4.2%	20.8%	16.7%	16.7%

Table 9.11. Distribution of Responses for Comparison Normal Academic Class (NHPS)

General Views		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q1	*pre*	0.0%	0.0%	0.0%	12.8%	7.7%	23.1%	35.9%	0.0%	20.5%
	post	5.0%	10.0%	0.0%	5.0%	10.0%	15.0%	25.0%	10.0%	20.0%
Q5	*pre*	5.1%	2.6%	10.3%	12.8%	10.3%	41.0%	12.8%	2.6%	2.6%
	post	0.0%	10.0%	5.0%	15.0%	5.0%	50.0%	10.0%	5.0%	0.0%
Q9	*pre*	2.6%	0.0%	0.0%	7.7%	12.8%	23.1%	20.5%	10.3%	23.1%
	post	5.0%	5.0%	0.0%	5.0%	15.0%	25.0%	5.0%	20.0%	20.0%
Q13	*pre*	15.4%	5.1%	23.1%	10.3%	25.6%	10.3%	7.7%	0.0%	2.6%
	post	21.1%	10.5%	26.3%	0.0%	15.8%	21.1%	5.3%	0.0%	0.0%
Q16.	*pre*	5.1%	5.1%	7.7%	17.9%	23.1%	17.9%	7.7%	12.8%	2.6%
	post	5.0%	0.0%	10.0%	15.0%	10.0%	20.0%	20.0%	5.0%	15.0%
Q19	*pre*	28.2%	2.6%	17.9%	17.9%	12.8%	12.8%	5.1%	0.0%	2.6%
	post	10.0%	5.0%	45.0%	15.0%	15.0%	10.0%	0.0%	0.0%	0.0%
Anxiety		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q2	*pre*	10.3%	2.6%	10.3%	12.8%	35.9%	20.5%	2.6%	0.0%	5.1%
	post	0.0%	0.0%	5.0%	40.0%	20.0%	25.0%	10.0%	0.0%	0.0%
Q6	*pre*	0.0%	0.0%	5.3%	21.1%	18.4%	15.8%	23.7%	0.0%	15.8%
	post	0.0%	0.0%	0.0%	15.0%	25.0%	25.0%	15.0%	10.0%	10.0%
Q10	*pre*	10.3%	2.6%	15.4%	10.3%	17.9%	38.5%	2.6%	0.0%	2.6%
	post	0.0%	15.0%	10.0%	10.0%	25.0%	25.0%	5.0%	10.0%	0.0%
Q14	*pre*	2.6%	5.1%	10.3%	20.5%	15.4%	20.5%	12.8%	7.7%	5.1%
	post	0.0%	10.5%	15.8%	10.5%	5.3%	36.8%	15.8%	5.3%	0.0%

Q17	*pre*	12.8%	2.6%	10.3%	7.7%	28.2%	12.8%	12.8%	7.7%	5.1%
	post	0.0%	5.0%	10.0%	20.0%	20.0%	25.0%	20.0%	0.0%	0.0%
Q20	*pre*	7.7%	0.0%	7.7%	28.2%	23.1%	10.3%	12.8%	10.3%	0.0%
	post	0.0%	0.0%	15.0%	30.0%	15.0%	20.0%	10.0%	5.0%	5.0%
Performance		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q3	*pre*	2.6%	0.0%	2.6%	2.6%	12.8%	25.6%	23.1%	17.9%	12.8%
	post	0.0%	0.0%	0.0%	10.0%	20.0%	50.0%	5.0%	10.0%	5.0%
Q7	*pre*	0.0%	5.1%	10.3%	7.7%	35.9%	28.2%	5.1%	5.1%	2.6%
	post	0.0%	0.0%	5.0%	15.0%	20.0%	45.0%	10.0%	5.0%	0.0%
Q11	*pre*	2.6%	10.3%	15.4%	5.1%	17.9%	17.9%	12.8%	12.8%	5.1%
	post	5.0%	0.0%	15.0%	0.0%	5.0%	35.0%	35.0%	0.0%	5.0%
Q15	*pre*	7.7%	2.6%	23.1%	10.3%	17.9%	17.9%	10.3%	2.6%	7.7%
	post	0.0%	15.8%	26.3%	21.1%	5.3%	21.1%	10.5%	0.0%	0.0%
Q21	*pre*	5.1%	5.1%	7.7%	23.1%	12.8%	20.5%	12.8%	7.7%	5.1%
	post	0.0%	5.0%	0.0%	20.0%	35.0%	15.0%	20.0%	5.0%	0.0%
Q22	*pre*	10.3%	2.6%	7.7%	12.8%	25.6%	10.3%	5.1%	7.7%	17.9%
	post	0.0%	15.0%	10.0%	20.0%	25.0%	20.0%	0.0%	0.0%	10.0%
Belief		DT	DAL	D	DLL	N	ALL	A	AAL	AT
Q4	*pre*	0.0%	0.0%	0.0%	10.3%	0.0%	7.7%	25.6%	5.1%	51.3%
	post	0.0%	0.0%	0.0%	0.0%	5.0%	5.0%	35.0%	15.0%	40.0%
Q8	*pre*	0.0%	2.6%	2.6%	7.7%	10.3%	10.3%	20.5%	23.1%	23.1%
	post	0.0%	0.0%	0.0%	0.0%	10.0%	10.0%	25.0%	30.0%	25.0%
Q12	*pre*	48.7%	15.4%	17.9%	5.1%	7.7%	2.6%	0.0%	2.6%	0.0%
	post	45.0%	0.0%	30.0%	10.0%	10.0%	5.0%	0.0%	0.0%	0.0%
Q18	*pre*	7.7%	0.0%	2.6%	0.0%	41.0%	23.1%	5.1%	5.1%	15.4%
	post	0.0%	0.0%	0.0%	5.0%	15.0%	25.0%	20.0%	10.0%	25.0%

In a similar way, before the start of the study and after 18 months form the start of the intervention, comparisons of students' attitudes toward mathematics and mathematics learning between the experimental and the corresponding comparison Normal Academic classes were examined. At the beginning of the study, there were significant differences in students' responses to question items Q8B (U [39, 39] = 535.00, $p < .05$) and Q18B (U [39, 39] = 566.50, $p < .05$). Both these items were pertaining to students' *'beliefs about the usefulness of mathematics'* and significantly more students in the experimental class believed that '*mathematics is useful'* as well as they would '*use mathematics a lot as an adult'*. However in the post-results, only item Q3P ((U [24, 20] = 121.00, $p < .005$) showed significant difference in students' responses between the experimental and comparison class. More specifically, more students in the experimental class believed that they sure could *'learn mathematics well'* compared to those in the comparison class. Students exposed to more writing and oral tasks with the encouragement of the teacher might have contributed to the above phenomenon.

Pre- and post-questionnaire surveys: About students' experiences with new assessment strategies

Nine items were constructed in this part B of the pre-questionnaire survey that all students from the experimental and comparison classes received, and the same nine items in the part B of the post-questionnaire survey that only the students from the comparison classes received. The nine items were constructed with the intention of measuring students' past experiences with some of the new learning strategies in the mathematics classroom. The other objective was to measure if the comparison classes were intact in their exposure to the new learning strategies during the intervention period of the study compared with the corresponding experimental classes. The nine items were specifically:

Q23a: my math teacher had asked me to write down the reasons for my math answers
Q23b: my math teacher had asked me to explain mathematics to the whole class.
Q23c: my math teacher had asked me to write down my feelings about mathematics.
Q23d: my math teacher had asked me to explain math ideas in writing.
Q24: my math teacher encouraged me to solve math questions in different ways.
Q25: my math teacher asked me to make up math questions by myself.
Q26: how often did your math teacher ask you think about the reason for your solving math problems?
Q27: how many math questions did your teacher ask you to do have more than 1 correct answer?
Q28: how many math questions did your teacher ask you to do have nothing to do with real life situations?

1. High-performing school

This sub-study focused on the use of communication tasks in mathematics classrooms. The items in part B of the pre-questionnaire survey that were constructed to measure students' experiences with some of the new strategies in mathematics learning and items that were examined in this study thought to be relevant were question items Q23a, Q23b, Q23c and Q23d. In this section, the report would first describe the results obtained from the Express stream classes followed by the Normal Academic stream classes.

Express stream. Table 9.12 shows the percentages of the Express students' responses to the individual items in the experimental and comparison classes.

Table 9.12. Distribution of express stream students' responses to Part B of pre-questionnaire surveys (HPS)

Question/Students		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*Experimental*	52.5%	12.5%	2.5%	0%	20%	12.5%
	Comparison	32.4%	0%	8.1%	0%	35.1%	24.3%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*Experimental*	17.5%	10%	5%	2.5%	40%	25%
	Comparison	26.3%	7.9%	13.2%	0%	39.5%	13.2%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*Experimental*	0%	0%	2.5%	2.5%	22.5%	72.5%
	Comparison	0%	0%	0%	0%	21.6%	78.4
Q23d: my math teacher had asked me to explain math ideas in writing.	*Experimental*	30%	25%	7.5%	0%	12.5%	25%
	Comparison	18.9%	5.4%	8.1%	0%	37.8%	29.7%

Using Mann-Whitney U test, the data revealed that there was a significant difference in the students' responses to item Q23a and Q23d between the two classes with U [40, 37] = 534.00, p = 0.017 and U [40, 37] = 563.50, p = 0.037 respectively. That is to say, at the beginning of the study, there were a higher percentage of students (52.5% saying that *almost everyday*) in the experimental class that had experience of '*writing down the reasons for my math answers'* compared to the comparison class that has only a percentage of 32.4%. Similarly, at least 55% of the students in the experimental class believed that they were more exposed to the idea of '*explaining math ideas in writing*' than the students (at least 24.3%) in the comparison class.

In order to detect that students in the comparison Express class were in tact in their experiences during the intervention period, these students were asked again to response to the four aforementioned question items Q23a, Q23b, Q23c and Q23d in the post-questionnaire survey. The study also made a comparison in the percentages of students' responses to these items in the pre- and post-questionnaire survey. The data in Table 9.13 revealed that students' responses were generally consistent between their responses in the pre- and post-questionnaire surveys;, that is, the students in the comparison class were in general kept intact in terms of their experiences about '*writing down reasons for my math answers*', '*explaining math to whole class*', '*writing down my feelings about math*' and '*explaining math ideas in writing*' throughout the intervention period.

Table 9.13. Distribution of comparison express stream students' responses to Part B of pre- and post-questionnaire surveys (HPS)

Question/Survey		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*pre*	32.4%	0%	8.1%	0%	35.1%	24.3%
	post	10.3%	23.1%	7.7%	2.6%	38.5%	17.9%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*pre*	26.3%	7.9%	13.2%	0%	39.5%	13.2%
	post	28.2%	12.8%	2.6%	2.6%	38.5%	15.4%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*pre*	0%	0%	0%	0%	21.6%	78.4
	post	.0%	7.7%	.0%	.0%	41.0%	51.3%
Q23d: my math teacher had asked me to explain math ideas in writing.	*pre*	18.9%	5.4%	8.1%	0%	37.8%	29.7%
	post	7.7%	15.4%	.0%	5.1%	43.6%	28.2%

Normal Academic stream. The percentages of students responding to items Q23a-Q23d in the pre-questionnaire survey for both the experimental and comparison Normal Academic classes were given in Table 9.14. The data revealed that there was a significant difference in the students' responses to item Q23b (U [40, 41] = 590.00, p = 0.015), with about 42.5% of the students in the experimental class claiming that in their past experience, their mathematics teachers 'rarely' or 'never' 'had asked me to explain mathematics to the whole class' compared to 63.4% of the comparison students making this claim. Other than this item, the students in both the experimental and comparison classes seemed to have equal experiences in doing new learning strategies that focused on communication at the beginning of the study. In fact, more than 40% of the students in both classes were either 'rarely' or 'never' been exposed to 'writing down reasons for my math answers', 'writing down my feelings about mathematics' or 'explaining math ideas in writing'.

Table 9.14. Distribution of both experimental and comparison normal academic students' responses to Part B of pre-questionnaire survey (HPS)

Question/Students		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*Experimental*	45.0%	7.5%	5.0%	0%	20.0%	22.5%
	Comparison	24.4%	9.8%	7.3%	2.4%	26.8%	29.3%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*Experimental*	37.5%	10.0%	7.5%	2.5%	25.0%	17.5%
	Comparison	19.5%	12.2%	4.9%	0%	24.4%	39.0%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*Experimental*	5.0%	10.0%	2.5%	2.5%	17.5%	62.5%
	Comparison	0%	0%	4.9%	7.3%	41.5%	46.3%
Q23d: my math teacher had asked me to explain math ideas in writing.	*Experimental*	32.5%	10.0%	12.5%	2.5%	12.5%	30.0%
	Comparison	12.2%	22.0%	9.8%	2.4%	26.8%	26.8%

Similarly, in order to detect whether students in the comparison class were in tact in their experiences dealing with new learning strategies that focused in communication, these students were asked again to response to the four aforementioned items Q23a-Q23d in the post-questionnaire survey. The data in Table 9.15 revealed that students were consistent in their experiences (with over 50% of the students responding to both in the pre- and post-questionnaire surveys either '*rarely*' or '*never*') in '*writing*' or '*explaining*' mathematics in classes.

Table 9.15. Distribution of comparison normal academic students' responses to Part B of pre- and post-questionnaire surveys (HPS)

Question/Survey		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*pre*	24.4%	9.8%	7.3%	2.4%	26.8%	29.3%
	post	28.6%	7.1%	10.7%	3.6%	39.3%	10.7%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*pre*	19.5%	12.2%	4.9%	0%	24.4%	39.0%
	post	39.3%	7.1%	.0%	.0%	35.7%	17.9%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*pre*	0%	0%	4.9%	7.3%	41.5%	46.3%
	post	3.6%	.0%	3.6%	3.6%	39.3%	50.0%
Q23d: my math teacher had asked me to explain math ideas in writing.	*pre*	12.2%	22.0%	9.8%	2.4%	26.8%	26.8%
	post	7.1%	7.1%	3.6%	10.7%	46.4%	25.0%

2. Non-high-performing school

In a similarly way, the report in this section would also first describe the results obtained from the Express stream classes followed by classes from the Normal Academic stream.

Express stream. Recall that the question items in part B of the pre-questionnaire survey were intended to measure students' experiences with some of the new strategies in mathematics learning and the items that were examined in this study specifically thought to be of relevance to communication were items Q23a, Q23b, Q23c and Q23d. At the beginning, the study analyzed the results between the experimental and the corresponding comparison class. Using Mann-Whitney tests, the data in Table 9.16 revealed that there were no significant differences in the students' responses to the abovementioned items. That was to say, at the beginning of the study, both the experimental and comparison students seemed to have equivalent exposure to these new learning strategies.

Table 9.16. Distribution of express stream students' responses to Part B of pre-questionnaire surveys (NHPS)

Question/Students		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*Experimental*	36.8%	7.9%	15.8%	2.6%	23.7%	13.2%
	Comparison	39.5%	7.9%	2.6%	.0%	34.2%	15.8%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*Experimental*	15.8%	10.5%	13.2%	.0%	31.6%	28.9%
	Comparison	26.3%	7.9%	7.9%	2.6%	26.3%	28.9%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*Experimental*	.0%	.0%	10.8%	2.7%	16.2%	70.3%
	Comparison	.0%	5.3%	.0%	2.6%	28.9%	63.2%
Q23d: my math teacher had asked me to explain math ideas in writing.	*Experimental*	23.7%	18.4%	2.6%	2.6%	21.1%	31.6%
	Comparison	36.8%	7.9%	5.3%	2.6%	21.1%	26.3%

In order that students in the comparison Express class were intact in their experiences during the intervention period, these students responded again to the same four aforementioned items Q23a-Q23d in the post-questionnaire survey. The percentages of students' responses in the pre- and post-questionnaire were recorded as shown in Table 9.17. The data revealed that the students responses seemed to have become less exposed to the new strategies of learning; for examples, for items Q23b-Q23d, there were increase in percentages of students responding to either '*rarely'* or '*never'* in the post-questionnaire compared to the results in the pre-questionnaire. Except for item Q23a, students seemed to have an increased in opportunity for them to *'write down reasons for my math answers'* compared to the beginning of the study.

Table 9.17. Distribution of comparison express stream students' responses to Part B of pre- and post-questionnaire surveys (NHPS)

Question/Survey		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*pre*	39.5%	7.9%	2.6%	0.0%	34.2%	15.8%
	post	36.4%	22.7%	9.1%	4.5%	22.7%	4.5%
Q23b: my math teacher had asked me to explain math to the whole class.	*pre*	26.3%	7.9%	7.9%	2.6%	26.3%	28.9%
	post	13.6%	0.0%	0.0%	0.0%	36.4%	50.0%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*pre*	0.0%	5.3%	0.0%	2.6%	28.9%	63.2%
	post	0.0%	0.0%	0.0%	0.0%	18.2%	81.8%
Q23d: my math teacher had asked me to explain math ideas in writing.	*pre*	36.8%	7.9%	5.3%	2.6%	21.1%	26.3%
	post	4.5%	4.5%	0.0%	4.5%	22.7%	63.6%

Normal Academic stream. The percentages of students responding to items Q23a-23d in the pre-questionnaire survey for the experimental and comparison Normal Academic classes were reflected in Table 9.18. The data revealed that there were no significant differences in the responses between the two classes. In other word, both the classes seemed to have equivalent experiences to the new strategies in mathematics learning that were explicitly stated in the above four mentioned items.

Table 9.18. Distribution of normal academic students' responses to Part B of pre- and post-questionnaire surveys (NHPS)

Question/Students		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*Experimental*	20.5%	20.5%	2.6%	.0%	33.3%	23.1%
	Comparison	33.3%	10.3%	5.1%	5.1%	15.4%	30.8%
Q23b: my math teacher had asked me to explain mathematics to the whole class.	*Experimental*	23.1%	23.1%	7.7%	2.6%	25.6%	17.9%
	Comparison	41.0%	2.6%	2.6%	.0%	28.2%	25.6%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*Experimental*	.0%	7.7%	5.1%	10.3%	23.1%	53.8%
	Comparison	.0%	2.6%	2.6%	2.6%	25.6%	66.7%
Q23d: my math teacher had asked me to explain math ideas in writing.	*Experimental*	17.9%	7.7%	7.7%	2.6%	30.8%	33.3%
	Comparison	23.1%	17.9%	7.7%	2.6%	15.4%	33.3%

In addition, to detect if the students in the comparison class were intact in their experiences to the four aforementioned learning strategies, these students responded to the items both in the pre- and post-questionnaire surveys. The data in Table 9.19 revealed that students appeared to have increased in their frequency in '*writing down the reasons for my math answers'* from 10.3% doing it'*2 or 3 times a week'* to 20%; and for item Q23d, students seemed to have also increased in their frequency in '*explaining math ideas in writing'* from total percentage of 10.3% doing it '*once a week or two weeks'/'once a month'* to 25%.

Table 9.19. Distribution of comparison normal academic students' responses to Part B of pre- and post-questionnaire surveys (NHPS)

Question/Survey		Almost every day	2 or 3 times a week	Once a week or two weeks	Once a month	Rarely	Never
Q23a: my math teacher had asked me to write down the reasons for my math answers	*pre*	33.3%	10.3%	5.1%	5.1%	15.4%	30.8%
	post	35.0%	20.0%	10.0%	0.0%	25.0%	10.0%
Q23b: my math teacher had asked me to explain math to the whole class.	*pre*	41.0%	2.6%	2.6%	0.0%	28.2%	25.6%
	post	25.0%	10.0%	10.0%	5.0%	30.0%	20.0%
Q23c: my math teacher had asked me to write down my feelings about mathematics.	*pre*	0.0%	2.6%	2.6%	2.6%	25.6%	66.7%
	post	5.0%	10.0%	5.0%	0.0%	15.0%	65.0%
Q23d: my math teacher had asked me to explain math ideas in writing.	*pre*	23.1%	17.9%	7.7%	2.6%	15.4%	33.3%
	post	30.0%	10.0%	10.0%	15.0%	20.0%	15.0%

Post-questionnaire survey: About students' attitudes toward communication tasks
Unlike the students in the comparison classes who responded to the same nine question items in Part B of the pre- and post-questionnaire surveys, the students in the experimental classes received a set of different question items in Part B of the post-questionnaire survey. A total of 31 items were constructed to measure students' general (a) perceptions about their own ability to perform, (b) beliefs in the usefulness, and (c) acceptance of communication tasks. Since communication tasks in this study comprised of the use of journal writing and oral presentation, 15 items out of the 31 items were about journal writing while the rest (16 items) were about oral presentation.

1. High-performing school

As said earlier, there were only two experimental classes receiving intervention in each school in this sub-project. In this section, the report would first describe the results obtained from the experimental Express stream class followed by the results obtained from the experimental Normal Academic class.

Express stream. Table 9.20 shows the percentages of the experimental Express students responding to the 31 items mentioned above (DT: Disagree totally; DAL: Disagree a lot; D: Disagree; DLL: Disagree a little; N: Neither disagree nor agree; ALL: Agree a little; A: Agree; AAL: Agree a lot; AT: Agree totally. The same notations apply to the tables below).

Table 9.20. Distribution of experimental express stream students' response to Part B of post-questionnaire survey (HPS)

Item	DT	DAL	D	DLL	N	ALL	A	AAL	AT
General Beliefs about Journal Writing									
Q25: Writing mathematics journals helps me to learn mathematics.	10.0%	5.0%	22.5%	10.0%	15.0%	12.5%	20.0%	2.5%	2.5%
Qn27: Writing mathematics journals helps me to be more aware of my understanding of mathematics.	5.0%	2.5%	15.0%	7.5%	15.0%	20.0%	20.0%	7.5%	7.5%
Qn28: Writing mathematics journals is an important skill in mathematics learning.	5.0%	10.0%	15.0%	20.0%	17.5%	5.0%	17.5%	7.5%	2.5%

Qn32: Writing mathematics journals makes me think broader and deeper about mathematics.	5.0%	2.5%	12.5%	17.5%	15.0%	15.0%	25.0%	5.0%	2.5%
Qn36: Writing mathematics journals makes me learn mathematics better.	7.5%	5.0%	17.5%	10.0%	15.0%	17.5%	20.0%	5.0%	2.5%
Qn37: Writing mathematics journals is a waste of time.	5.0%	7.5%	7.5%	15.0%	17.5%	30.0%	10.0%	.0%	7.5%
Qn29: I am able to express about my feeling toward mathematics through writing mathematics journals.	12.5%	.0%	25.0%	5.0%	20.0%	12.5%	20.0%	2.5%	2.5%
Qn30: I am able to tell others about my understanding of mathematics through writing mathematics journals.	7.5%	.0%	12.5%	10.0%	10.0%	15.0%	32.5%	5.0%	7.5%
Perceptions about own Ability to do journal writing									
Q24: Writing mathematics journals is easy to me.	.0%	10.0%	22.5%	10.0%	12.5%	20.0%	20.0%	2.5%	2.5%
Qn26: I am not afraid of writing mathematics journals.	.0%	2.5%	10.0%	2.5%	17.5%	27.5%	27.5%	10.0%	2.5%
Qn31: I don't know how to get started when I am writing mathematics journals.	.0%	7.5%	15.0%	15.0%	15.0%	20.0%	15.0%	5.0%	7.5%
Qn33: When I am writing mathematics journals, I know what I am expected to write.	.0%	.0%	10.0%	17.5%	12.5%	17.5%	25.0%	12.5%	5.0%
Qn34: I can write my mathematics journals well.	7.5%	5.0%	22.5%	10.0%	25.0%	7.5%	20.0%	.0%	2.5%
General Acceptance towards journal writing									
Q23: I like to write mathematics journals.	15.0%	12.5%	20.0%	10.0%	15.0%	12.5%	7.5%	7.5%	.0%
Qn35: I would like to have more mathematics journal writing for my mathematics lessons	25.0%	15.0%	17.5%	5.0%	22.5%	2.5%	7.5%	.0%	5.0%
General Beliefs about oral presentation									
Qn40: Doing mathematics oral presentation helps me to learn mathematics.	10.0%	7.5%	10.0%	15.0%	17.5%	15.0%	15.0%	5.0%	5.0%
Qn42: Doing mathematics oral presentation helps me to be more aware of my understanding of mathematics.	7.5%	5.0%	10.0%	10.0%	22.5%	20.0%	20.0%	2.5%	2.5%
Qn43: Oral presentation skill is important in mathematics learning.	12.5%	7.5%	12.5%	12.5%	22.5%	10.0%	17.5%	.0%	5.0%
Qn47: Doing mathematics oral presentation makes me think broader and deeper about mathematics.	2.5%	2.5%	15.0%	15.0%	22.5%	17.5%	17.5%	2.5%	5.0%
Qn49: Listening to other classmates' oral presentation is helpful for me in learning mathematics.	2.5%	.0%	2.5%	10.0%	12.5%	20.0%	32.5%	7.5%	12.5%
Qn52: Doing mathematics oral presentation makes me learn mathematics better.	7.5%	15.0%	7.5%	5.0%	20.0%	20.0%	17.5%	5.0%	2.5%
Qn53: Doing mathematics oral	10.0%	7.5%	7.5%	2.5%	37.5%	5.0%	12.5%	5.0%	12.5%

presentation is a waste of time.									
Qn44: I am able to express about my feeling through mathematics oral presentation.	7.5%	5.0%	20.0%	15.0%	20.0%	12.5%	12.5%	5.0%	2.5%
Qn45: I am able to tell others about my understanding of mathematics through mathematics oral presentation.	7.5%	.0%	20.0%	10.0%	17.5%	20.0%	17.5%	2.5%	5.0%
Perceptions about own Ability to do oral presentation									
Qn39: Doing mathematics oral presentation is easy to me.	15.0%	10.0%	17.5%	15.0%	20.0%	12.5%	5.0%	5.0%	.0%
Qn41: I am not afraid of doing mathematics oral presentation.	15.0%	2.5%	15.0%	17.5%	12.5%	7.5%	25.0%	5.0%	.0%
Qn46: I don't know how to get started when I am doing mathematics oral presentation.	2.5%	.0%	15.0%	7.5%	30.0%	22.5%	7.5%	5.0%	10.0%
Qn48: I feel lost when I am doing mathematics oral presentation.	2.5%	5.0%	10.0%	15.0%	22.5%	20.0%	15.0%	7.5%	2.5%
Qn50: I can do mathematics oral presentation well.	10.0%	7.5%	20.0%	15.0%	25.0%	12.5%	7.5%	2.5%	.0%
General Acceptance towards oral presentation									
Qn38: I like to do mathematics oral presentation during mathematics lessons	25.0%	12.5%	12.5%	10.0%	20.0%	7.5%	10.0%	.0%	2.5%
Qn51: I would like to have more mathematics oral presentations for my mathematics lessons.	25.0%	12.5%	7.5%	12.5%	22.5%	7.5%	5.0%	2.5%	2.5%

The above data revealed that students responded positively to 5 items pertaining to journal writing and 4 items pertaining to oral presentation. Although students did not generally have positive acceptance towards the use of journal writing (Q23 & Q35) and oral presentation (Q38 & Q51), more than 50% of them did believed that *'writing mathematics journals helps me to be more aware of my understanding of mathematics'* (Q27), and 47.5% have the same opinion that '*writing mathematics journals makes me think broader and deeper about mathematics'* (Q32).

Similarly, these students also had positive beliefs about the usefulness of doing oral presentation. For instance in Q42, 45% of the students believed that *'doing mathematics oral presentation helps me to be more aware of my understanding of mathematics* (vs. 32.5% who disagreed), 42.5% of the students believed that '*doing mathematics oral presentation makes me think broader and deeper about mathematics'* for Q47 (vs. 35% who disagreed), and more than 70% of the students also had the same opinion that '*listening to other classmates' oral presentation is helpful for me in learning mathematics'* (Q49).

The above results were not unexpected because consistently it was revealed in the interviews with students that they were generally still not very confident in their performance in writing and oral. They had expressed that more practice were needed for them to improve and make progress. However, the students did believe that ultimately, the practice of more proper and concise writing, and opportunities to speak up in class would make them learn mathematics better.

Normal Academic stream. Table 9.21 shows the percentages of the experimental Normal Academic students responding to the 31 aforementioned items in the above. The data for the experimental Normal Academic students interestingly also revealed somehow similarly results as to the experimental Express students' responses to these items. The

Normal Academic students did not show very positive responses to items corresponding to their acceptance towards the use of journal writing and oral presentation.

However, the data did reveal that most students generally believed in the usefulness of these new communication tasks. For instance, 59% of the students 'agreed' that *'writing mathematics journals helps me to be more aware of my understanding of mathematics',* 43.6% 'agreed' that '*writing mathematics journals is an important skill in mathematics learning'* (vs. 25.6% who 'disagreed'), 51.2% believed that '*writing mathematics journals makes me think broader and deeper about mathematics'* , and more than 60% of the students 'agreed' that *'listening to other classmates' oral presentation is helpful for me in learning mathematics'.* In the interviews with students, they had similar feelings as their counterparts in the Express class. They believed with more practice, they would ultimately be able to do better in writing and speaking.

Table 9.21. Distribution of experimental normal academic students' responses to Part B of post-questionnaire survey (HPS)

Item	DT	DAL	D	DLL	N	ALL	A	AAL	AT
General beliefs about journal writing									
Q25: Writing mathematics journals helps me to learn mathematics.	20.5%	5.1%	7.7%	7.7%	17.9%	23.1%	7.7%	5.1%	5.1%
Qn27: Writing mathematics journals helps me to be more aware of my understanding of mathematics.	7.7%	7.7%	2.6%	2.6%	20.5%	25.6%	15.4%	7.7%	10.3%
Qn28: Writing mathematics journals is an important skill in mathematics learning.	12.8%	5.1%	5.1%	2.6%	30.8%	17.9%	15.4%	2.6%	7.7%
Qn32: Writing mathematics journals makes me think broader and deeper about mathematics.	7.7%	.0%	5.1%	5.1%	30.8%	20.5%	20.5%	5.1%	5.1%
Qn36: Writing mathematics journals makes me learn mathematics better.	17.9%	2.6%	7.7%	10.3%	15.4%	20.5%	7.7%	12.8%	5.1%
Qn37: Writing mathematics journals is a waste of time.	17.9%	5.1%	15.4%	12.8%	20.5%	10.3%	5.1%	2.6%	10.3%
Qn29: I am able to express about my feeling toward mathematics through writing mathematics journals.	10.3%	5.1%	7.7%	2.6%	12.8%	30.8%	12.8%	7.7%	10.3%
Qn30: I am able to tell others about my understanding of mathematics through writing mathematics journals.	10.3%	5.1%	5.1%	2.6%	20.5%	28.2%	10.3%	10.3%	7.7%
Perceptions about own ability to do journal writing									
Q24: Writing mathematics journals is easy to me.	10.3%	7.7%	15.4%	10.3%	20.5%	17.9%	7.7%	2.6%	7.7%
Qn26: I am not afraid of writing mathematics journals.	10.5%	2.6%	7.9%	5.3%	15.8%	18.4%	23.7%	5.3%	10.5%
Qn31: I don't know how to get started when I am writing mathematics journals.	12.8%	.0%	10.3%	5.1%	20.5%	20.5%	15.4%	7.7%	7.7%
Qn33: When I am writing mathematics journals, I know what I am expected to write.	15.4%	.0%	7.7%	12.8%	23.1%	17.9%	12.8%	2.6%	7.7%
Qn34: I can write my mathematics journals well.	17.9%	.0%	10.3%	17.9%	20.5%	20.5%	5.1%	.0%	7.7%

General acceptance towards journal writing									
Q23: I like to write mathematics journals.	23.1%	2.6%	20.5%	5.1%	12.8%	12.8%	17.9%	2.6%	2.6%
Qn35: I would like to have more mathematics journal writing for my mathematics lessons	25.6%	2.6%	5.1%	10.3%	23.1%	20.5%	7.7%	2.6%	2.6%
General beliefs about oral presentation									
Qn40: Doing mathematics oral presentation helps me to learn math.	23.1%	5.1%	17.9%	2.6%	28.2%	10.3%	10.3%	2.6%	.0%
Qn42: Doing mathematics oral presentation helps me to be more aware of my understanding of mathematics.	23.1%	2.6%	10.3%	7.7%	23.1%	17.9%	7.7%	5.1%	2.6%
Qn43: Oral presentation skill is important in mathematics learning.	17.9%	2.6%	10.3%	7.7%	25.6%	17.9%	7.7%	.0%	10.3%
Qn47: Doing mathematics oral presentation makes me think broader and deeper about mathematics.	17.9%	5.1%	7.7%	10.3%	23.1%	15.4%	12.8%	2.6%	5.1%
Qn49: Listening to other classmates' oral presentation is helpful for me in learning mathematics.	12.8%	2.6%	.0%	5.1%	10.3%	30.8%	17.9%	5.1%	15.4%
Qn52: Doing mathematics oral presentation makes me learn mathematics better.	25.6%	.0%	15.4%	5.1%	17.9%	20.5%	10.3%	2.6%	2.6%
Qn53: Doing mathematics oral presentation is a waste of time.	10.3%	2.6%	10.3%	10.3%	23.1%	12.8%	2.6%	5.1%	23.1%
Qn44: I am able to express about my feeling through mathematics oral presentation.	15.8%	5.3%	18.4%	5.3%	28.9%	10.5%	10.5%	2.6%	2.6%
Qn45: I am able to tell others about my understanding of mathematics through mathematics oral presentation.	17.9%	5.1%	15.4%	5.1%	25.6%	15.4%	10.3%	2.6%	2.6%
Perceptions about own ability to do oral presentation									
Qn39: Doing mathematics oral presentation is easy to me.	28.2%	10.3%	5.1%	17.9%	20.5%	10.3%	5.1%	.0%	2.6%
Qn41: I am not afraid of doing mathematics oral presentation.	15.4%	5.1%	10.3%	.0%	30.8%	7.7%	15.4%	.0%	15.4%
Qn46: I don't know how to get started when I am doing mathematics oral presentation.	7.7%	2.6%	12.8%	2.6%	17.9%	20.5%	10.3%	5.1%	20.5%
Qn48: I feel lost when I am doing mathematics oral presentation.	10.3%	2.6%	10.3%	5.1%	20.5%	15.4%	17.9%	5.1%	12.8%
Qn50: I can do mathematics oral presentation well.	25.6%	10.3%	12.8%	12.8%	17.9%	10.3%	7.7%	2.6%	.0%
General acceptance towards oral presentation									
Qn38: I like to do mathematics oral presentation during mathematics lessons	30.8%	7.7%	10.3%	17.9%	17.9%	5.1%	7.7%	.0%	2.6%
Qn51: I would like to have more mathematics oral presentations for my mathematics lessons.	38.5%	5.1%	10.3%	10.3%	20.5%	10.3%	5.1%	.0%	.0%

2. Non-high-performing school

Express stream. Table 9.22 provided the percentages of the experimental Express students responding to the 31 items that were intended to measure students' *general beliefs of the usefulness*, *perceptions of own ability to perform*, , and *general acceptance* toward journal writing and oral presentation tasks in mathematics classrooms.

Table 9.22. Distribution of experimental express stream students' response to Part B of post-questionnaire survey (NHPS)

Item	DT	DAL	D	DLL	N	ALL	A	AAL	AT
General beliefs about journal writing									
Q25: Writing mathematics journals helps me to learn mathematics.	14.3%	4.8%	4.8%	9.5%	28.6%	23.8%	9.5%	0.0%	4.8%
Qn27: Writing mathematics journals helps me to be more aware of my understanding of mathematics.	10.0%	5.0%	5.0%	10.0%	10.0%	25.0%	25.0%	0.0%	10.0%
Qn28: Writing mathematics journals is an important skill in mathematics learning.	4.8%	0.0%	14.3%	19.0%	28.6%	14.3%	14.3%	0.0%	4.8%
Qn32: Writing mathematics journals makes me think broader and deeper about mathematics.	9.5%	4.8%	4.8%	4.8%	28.6%	14.3%	19.0%	9.5%	4.8%
Qn36: Writing mathematics journals makes me learn mathematics better.	14.3%	4.8%	4.8%	9.5%	28.6%	14.3%	14.3%	4.8%	4.8%
Qn37: Writing mathematics journals is a waste of time.	14.3%	14.3%	14.3%	19.0%	9.5%	4.8%	4.8%	4.8%	14.3%
Qn29: I am able to express about my feeling toward mathematics through writing math journals.	4.8%	9.5%	4.8%	4.8%	19.0%	23.8%	14.3%	4.8%	14.3%
Qn30: I am able to tell others about my understanding of mathematics through writing mathematics journals.	0.0%	4.8%	14.3%	4.8%	23.8%	19.0%	19.0%	0.0%	14.3%
Perceptions about own ability to do journal writing									
Q24: Writing mathematics journals is easy to me.	4.8%	4.8%	9.5%	14.3%	19.0%	9.5%	9.5%	9.5%	19.0%
Qn26: I am not afraid of writing mathematics journals.	4.8%	0.0%	9.5%	4.8%	23.8%	28.6%	9.5%	4.8%	14.3%
Qn31: I don't know how to get started when I am writing mathematics journals.	9.5%	0.0%	4.8%	9.5%	19.0%	33.3%	14.3%	4.8%	4.8%
Qn33: When I am writing mathematics journals, I know what I am expected to write.	14.3%	4.8%	4.8%	14.3%	19.0%	28.6%	14.3%	0.0%	0.0%
Qn34: I can write my mathematics journals well.	9.5%	0.0%	19.0%	28.6%	19.0%	19.0%	4.8%	0.0%	0.0%
General acceptance towards journal writing									
Q23: I like to write mathematics journals.	19.0%	4.8%	0.0%	28.6%	23.8%	9.5%	4.8%	0.0%	9.5%
Qn35: I would like to have more mathematics journal writing for my mathematics lessons	23.8%	4.8%	23.8%	4.8%	23.8%	9.5%	4.8%	0.0%	4.8%
General beliefs about oral presentation									

Qn40: Doing mathematics oral presentation helps me to learn mathematics.	4.8%	4.8%	0.0%	9.5%	19.0%	28.6%	28.6%	0.0%	4.8%
Qn42: Doing mathematics oral presentation helps me to be more aware of my understanding of mathematics.	4.8%	4.8%	0.0%	0.0%	19.0%	38.1%	19.0%	4.8%	9.5%
Qn43: Oral presentation skill is important in mathematics learning.	4.8%	0.0%	9.5%	0.0%	23.8%	23.8%	23.8%	4.8%	9.5%
Qn47: Doing mathematics oral presentation makes me think broader and deeper about mathematics.	9.5%	0.0%	4.8%	0.0%	28.6%	33.3%	14.3%	4.8%	4.8%
Qn49: Listening to other classmates' oral presentation is helpful for me in learning mathematics.	0.0%	0.0%	0.0%	0.0%	14.3%	19.0%	42.9%	9.5%	14.3%
Qn52: Doing mathematics oral presentation makes me learn mathematics better.	4.8%	9.5%	4.8%	0.0%	28.6%	19.0%	19.0%	4.8%	9.5%
Qn53: Doing mathematics oral presentation is a waste of time.	28.6%	4.8%	9.5%	19.0%	19.0%	4.8%	4.8%	0.0%	9.5%
Qn44: I am able to express about my feeling through mathematics oral presentation.	9.5%	0.0%	23.8%	19.0%	23.8%	14.3%	9.5%	0.0%	0.0%
Qn45: I am able to tell others about my understanding of mathematics through mathematics oral presentation.	9.5%	0.0%	23.8%	4.8%	23.8%	23.8%	4.8%	9.5%	0.0%
Perceptions about own ability to do oral presentation									
Qn39: Doing mathematics oral presentation is easy to me.	23.8%	14.3%	14.3%	14.3%	23.8%	4.8%	4.8%	0.0%	0.0%
Qn41: I am not afraid of doing mathematics oral presentation.	19.0%	9.5%	9.5%	9.5%	33.3%	14.3%	0.0%	4.8%	0.0%
Qn46: I don't know how to get started when I am doing mathematics oral presentation.	0.0%	0.0%	4.8%	4.8%	23.8%	4.8%	19.0%	23.8%	19.0%
Qn48: I feel lost when I am doing mathematics oral presentation.	0.0%	4.8%	0.0%	9.5%	38.1%	9.5%	14.3%	9.5%	14.3%
Qn50: I can do mathematics oral presentation well.	23.8%	14.3%	4.8%	23.8%	23.8%	0.0%	4.8%	0.0%	4.8%
General acceptance towards oral presentation									
Qn38: I like to do mathematics oral presentation during mathematics lessons	23.8%	4.8%	14.3%	9.5%	14.3%	23.8%	4.8%	4.8%	0.0%
Qn51: I would like to have more mathematics oral presentations for my mathematics lessons.	19.0%	9.5%	14.3%	9.5%	33.3%	4.8%	0.0%	4.8%	4.8%

The data in Table 9.22 revealed that out of the 15 items that measured students' views about journal writing, students responded positively to seven of them; and out of 16 items that measured students' views toward oral presentation, seven items were responded positively, i.e. about 50% of the total items received positive responses. Although only 23.8% and 19.1% of the students would '*like to write mathematics journals'* (Q230 and '*like to have more mathematics journal writing for my mathematics lessons'*(Q35) respectively, there were a total of 61.9% of them who 'disbelieved' that '*writing* *mathematics journals is a waste of time'* (Q37). Moreover, 47.6% of the students 'agreed' that '*writing mathematics journals makes me think broader and deeper about mathematics'* (Q32) (compared to 23.9% who 'disagreed'); similarly 50% of the students also 'agreed' that '*writing mathematics journals helps me to be more aware of my understanding of mathematics'* (Q27) (vs. 30% who 'disagreed').

In addition, students were also found to be quite positive in their perceptions of own ability to do journal writing. For examples, in item Q24 '*writing mathematics journals is easy to me'* received 47.5% of students 'agreeing' to the statement (vs. 33.4% who 'disagreed'), and more than 50% 'agreed' that they were *'not afraid of writing mathematics journals'* (Q26).

Interestingly, the data above also showed that students had also responded quite negatively about their acceptance towards the use of oral presentation in the mathematics classroom. For instance, only 33.4% and 14.4% of the students 'agreed' to item Q38 '*I like to do mathematics oral presentation during mathematics lessons'* and item Q51 '*I would like to have more mathematics oral presentations for my mathematics lessons'* respectively. However, data revealed that more than 60% of the students believed that '*doing mathematics oral presentation is not a waste of time'* (Q53) and that '*oral presentation skill is important in mathematics learning'* (Q43).

In addition, there were high percentages of students who believed that doing oral presentations helped in their learning of mathematics. For example, 62% and 52.3% of them 'agreed' that '*doing mathematics oral presentation helps me to learn mathematics'* (Q40) and '*learn mathematics better'* (Q52)*;* 71.4% felt that '*doing mathematics oral presentation helps me to be more aware of my understanding of mathematics'* (Q42); 57.2% 'agreed' that '*doing mathematics oral presentation makes me think broader and deeper about mathematics'* (Q47); and over 80% 'agreed' that *'listening to other classmates' oral presentation is helpful for me in learning mathematics'* (49).

Compared to the students in the high-performing school, this group of students was more positive towards the use of communication tasks. This was believed to have contributed from the teacher herself. Although this teacher started the intervention only in the second year (Jan 2006 – May 2006), but she was very serious and consistent in integrating and implementing communication tasks into her instructions. At the same time, she had made known to her students that all the tasks, be it the normal conventional or communication tasks were equally important and students took all the work quite seriously.

Normal Academic stream. Table 9.23 provided the percentages of the experimental Normal Academic students responding to the 31 items of part B of the post-questionnaire survey.

The data showed that out of the 15 items measuring students' views toward the use of journal writing only five received positive responses, and out of 16 items measuring students' views toward the use of oral presentation only six received positive responses. Thus, less than 50% of the total number of items received positive feedback from this group of students.

Table 9.23 revealed that generally, students did not like to write mathematics journals or would like to have more mathematics journal tasks in their mathematics classroom. However, students did have positive views toward the usefulness of journal writing.

For examples, 62.5% of the students 'agreed' that *'writing mathematics journals helps me to be more aware of my understanding of mathematics'* (Q 27) and 54.2% also believed that '*writing mathematics journals makes me think broader and deeper about mathematics*' (Q32). Moreover, 37.5% 'agreed' (vs. 29.2% who 'disagreed') that '*writing* *mathematics journals is an important skill in mathematics learning*' (Q28). 75% of the students also responded that they were '*not afraid of writing mathematics journals'* (Q26).

In the same sense, this group of students also did not appear to have very positive acceptance towards oral presentation but did 'believed' (33.4% vs. 29.2% who 'disbelieved') that '*oral presentation skill is important in mathematics learning'* (Q43) and about 50% 'agreed' that '*doing mathematics oral presentation is not a waste of time'* (Q53). Moreover, the students had a general consensus that oral presentation helped in their learning of mathematics.

For examples, item Q42 '*doing mathematics oral presentation helps me to be more aware of my understanding of mathematics'* (Q42) had received 58.2% of students 'agreeing' to the item; 50.1% of the students 'agreed' that '*doing mathematics oral presentation makes me learn mathematics better'* (Q52); and 41.7% of the students 'believed' (vs. 25% who 'disbelieved') that '*doing mathematics oral presentation makes me think broader and deeper about mathematics*' (Q47).

Table 9.23. Distribution of experimental normal academic students' responses to Part B of post-questionnaire survey (HPS)

Item	DT	DAL	D	DLL	N	ALL	A	AAL	AT
General beliefs about journal writing									
Q25: Writing mathematics journals helps me to learn mathematics.	12.5%	0%	12.5%	12.5%	33.3%	0%	25%	4.2%	0%
Qn27: Writing mathematics journals helps me to be more aware of my understanding of mathematics.	12.5%	0.0%	0.0%	16.7%	8.3%	20.8%	25.0%	12.5%	4.2%
Qn28: Writing mathematics journals is an important skill in mathematics learning.	12.5%	0.0%	0.0%	16.7%	33.3%	20.8%	8.3%	4.2%	4.2%
Qn32: Writing mathematics journals makes me think broader and deeper about mathematics.	12.5%	0.0%	4.2%	12.5%	16.7%	16.7%	25.0%	8.3%	4.2%
Qn36: Writing mathematics journals makes me learn mathematics better.	12.5%	4.2%	16.7%	12.5%	25.0%	8.3%	12.5%	8.3%	0.0%
Qn37: Writing mathematics journals is a waste of time.	12.5%	0.0%	8.3%	20.8%	29.2%	4.2%	8.3%	0.0%	16.7%
Qn29: I am able to express about my feeling toward mathematics through writing mathematics journals.	12.5%	4.2%	12.5%	12.5%	16.7%	12.5%	20.8%	4.2%	4.2%
Qn30: I am able to tell others about my understanding of mathematics through writing mathematics journals.	12.5%	4.2%	0.0%	16.7%	8.3%	25.0%	12.5%	16.7%	4.2%
Perceptions about own ability to do journal writing									
Q24: Writing mathematics journals is easy to me.	16.7%	12.5%	12.5%	20.8%	29.2%	4.2%	4.2%	0.0%	0.0%
Qn26: I am not afraid of writing mathematics journals.	0.0%	0.0%	4.2%	8.3%	12.5%	25.0%	25.0%	8.3%	16.7%
Qn31: I don't know how to get started when I am writing mathematics journals.	0.0%	0.0%	4.2%	8.3%	12.5%	16.7%	25.0%	12.5%	20.8%
Qn33: When I am writing	12.5%	4.2%	8.3%	12.5%	20.8%	12.5%	20.8%	4.2%	4.2%

mathematics journals, I know what I am expected to write.									
Qn34: I can write my mathematics journals well.	20.8%	8.3%	20.8%	20.8%	16.7%	4.2%	8.3%	0.0%	0.0%
General acceptance towards journal writing									
Q23: I like to write mathematics journals.	20.8%	8.3%	16.7%	20.8%	16.7%	12.5%	4.2%	0.0%	0.0%
Qn35: I would like to have more mathematics journal writing for my mathematics lessons	29.2%	12.5%	12.5%	12.5%	25.0%	4.2%	4.2%	0.0%	0.0%
General beliefs about oral presentation									
Qn40: Doing mathematics oral presentation helps me to learn mathematics.	16.7%	0.0%	8.3%	8.3%	25.0%	16.7%	20.8%	0.0%	4.2%
Qn42: Doing mathematics oral presentation helps me to be more aware of my understanding of mathematics.	16.7%	0.0%	4.2%	4.2%	16.7%	20.8%	20.8%	8.3%	8.3%
Qn43: Oral presentation skill is important in mathematics learning.	16.7%	0.0%	4.2%	8.3%	37.5%	16.7%	12.5%	4.2%	0.0%
Qn47: Doing mathematics oral presentation makes me think broader and deeper about mathematics.	16.7%	0.0%	0.0%	8.3%	33.3%	12.5%	25.0%	4.2%	0.0%
Qn49: Listening to other classmates' oral presentation is helpful for me in learning mathematics.	12.5%	0.0%	0.0%	0.0%	8.3%	12.5%	37.5%	8.3%	20.8%
Qn52: Doing mathematics oral presentation makes me learn mathematics better.	16.7%	0.0%	0.0%	12.5%	20.8%	12.5%	29.2%	4.2%	4.2%
Qn53: Doing mathematics oral presentation is a waste of time.	12.5%	8.3%	20.8%	8.3%	33.3%	4.2%	0.0%	0.0%	12.5%
Qn44: I am able to express about my feeling through mathematics oral presentation.	16.7%	0.0%	8.3%	12.5%	33.3%	8.3%	16.7%	4.2%	0.0%
Qn45: I am able to tell others about my understanding of mathematics through mathematics oral presentation.	16.7%	4.2%	4.2%	8.3%	25%	8.3%	20.8%	12.5%	0%
Perceptions about own ability to do oral presentation									
Qn39: Doing mathematics oral presentation is easy to me.	20.8%	4.2%	20.8%	8.3%	20.8%	12.5%	4.2%	4.2%	4.2%
Qn41: I am not afraid of doing mathematics oral presentation.	25.0%	4.2%	12.5%	25.0%	4.2%	12.5%	12.5%	0.0%	4.2%
Qn46: I don't know how to get started when I am doing mathematics oral presentation.	8.3%	12.5%	12.5%	8.3%	25.0%	16.7%	4.2%	0.0%	12.5%
Qn48: I feel lost when I am doing mathematics oral presentation.	4.2%	4.2%	16.7%	12.5%	25.0%	16.7%	0.0%	4.2%	16.7%
Qn50: I can do mathematics oral presentation well.	25.0%	0.0%	12.5%	4.2%	41.7%	8.3%	0.0%	4.2%	4.2%
General acceptance towards oral presentation									

Qn38: I like to do mathematics oral presentation during mathematics lessons	20.8%	4.2%	20.8%	8.3%	33.3%	4.2%	0.0%	4.2%	4.2%
Qn51: I would like to have more mathematics oral presentations for my mathematics lessons.	29.2%	4.2%	20.8%	8.3%	20.8%	8.3%	4.2%	0.0%	4.2%

As mentioned earlier, this teacher was encouraging and supportive towards his students' learning. However, the teacher also believed that his students were generally very weak in their English language and he felt that this might contribute to his students' poor performance in both writing and oral, and thus students' negative attitudes toward communication tasks. In order to establish a more positive attitude, belief and liking towards the use of communication tasks, the teacher believed that his students would need more support, practice, and persuasion into doing the new tasks.

Summary

The 22 items of part A of the pre- and post-questionnaire surveys were intended to measure student's general attitudes toward mathematics and mathematics learning. The results revealed that all the students generally had more negative responses to the items in the post-questionnaire survey compared to the pre-questionnaire survey. That was to say, as students were promoted to secondary two from secondary one level, their general views and attitudes toward the subject mathematics and their own capability to perform the subject seemed to have more declined. Interestingly, the experimental classes in the non-high performing school had better attitudes compared to the experimental classes in the high-performing school. It was noted that the experimental classes in the non-high performing school had about 60% of the question items responded more positively in the post-questionnaire survey than the pre, and this was 24% more than the average percentage (36%) for the other classes (including the comparison classes).

The items in part B of the pre-questionnaire survey were intended to measure student's general past experiences with some of the new strategies in the mathematics classroom. The idea was to identify if the experimental and the corresponding comparison classes were equivalent or if there was any significant differences in their exposure to some of these new strategies before the start of the intervention in the experimental classes. Results revealed that generally, all the experimental classes compared to their corresponding comparison classes; they were equivalent in their experiences with some of the new strategies. In addition, the study also wanted to make sure that the comparison classes did not receive any of the new learning strategies in the mathematics classrooms throughout the whole intervention period and therefore they were asked to response to the same items of part B in the post-questionnaire survey again. The test for differences in their responses between the pre- and post-questionnaire surveys revealed that there were no significant changes. That was to say, the comparison classes were not influenced by the learning of the new strategies in their mathematics classrooms.

The 31 items in part B of the post-questionnaire survey for the experimental classes was intended to measure their general attitudes toward the new strategy tasks that they had worked on for about 18 months. Out of the 31 items, 15 items were questions about the use of journal writing while the 16 items were about the uses of oral presentation.

All the experimental classes except for the experimental Express class in the non-high performing school had less than 30% of the total items responded positively. Although students generally did not seem to accept more of the new tasks to be done in mathematics classrooms, they were rather positive about the beliefs of the benefits and usefulness of doing these new tasks. Students were by and large 'agreed' that doing journal writing and oral presentation "*helps me to learn mathematics", "helps me to be* *more aware of my understanding of mathematics",* and *"makes me think broader and deeper about mathematics".*

9.4.2 Students' achievement on school-based exams

This sub-study also examined students' achievement scores based on their school-based mathematics semester examinations. They were specifically the school's 2004 mid-year mathematics examination scores (M2004), school's 2004 final-year mathematics examination scores (F2004) and the school's 2005 mid-year mathematics examination scores (M2005). The intention was to examine and detect if there were any changes or impact on the experimental students' school mathematics results compared to their respective comparison classes after being exposed to about 18 months of communication tasks in their mathematics learning. This study also looked at the students' PSLE mathematics grades as a measurement to compare students' achievement level difference between the experimental and comparison classes at the beginning of the study.

The PSLE mathematics grades collected were not numeric and therefore chi-square test was used to detect if there were any differences between the experimental and comparison classes at the start of the study before the intervention was carried out in schools.

The results showed that there was no significant difference between the experimental and comparison classes in the Express stream (χ^2 [2, N = 78] = 0.117, p = .943) as well as in the Normal Academic stream (χ^2 [3, N = 81] = 2.562, p = .464). In term of the results, it appears that the experimental and comparison classes in the Express stream (as well as in the Normal Academic stream respectively) were equivalent at the start of the study based on the PSLE mathematics grades obtained.

High-performing school
Three sets of the school's mathematics scores which were namely M2004, F2004 and M2005 were collected from all the experimental and comparison classes. The mean scores were calculated and Table 9.24 revealed the mean scores of the respective mathematics examinations for each classes.

Table 9.24. Mean scores of school-based examinations (HPS)

Mean scores		M2004	F2004	M2005
Express	*Experimental*	68.28	63.68	63.90
	Comparison	62.37	59.13	52.59
Normal Academic	*Experimental*	59.45	53.55	61.19
	Comparison	59.76	52.59	56.46

Using the three sets of school's mathematics examination scores which were namely the M2004, F2004 and M2005, the differences in the mean scores between the experimental and the comparison classes of each of these examination scores were evaluated. Figure 9.1 showed the changes in the scores for both the Express and Normal Academic classes.

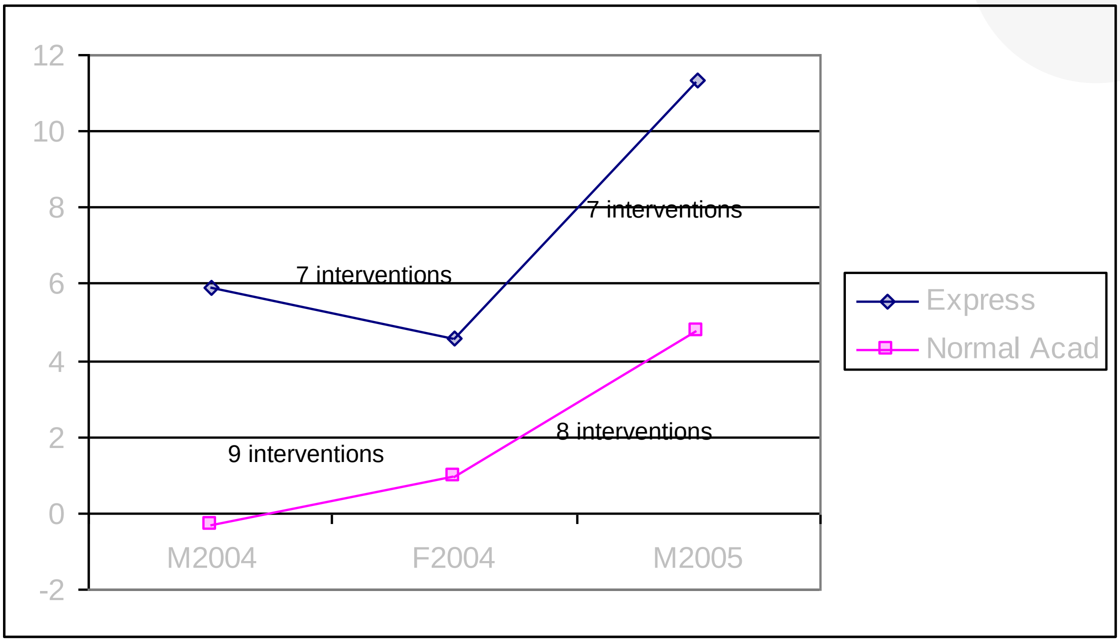


Figure 9.1. Differences in school-based exam mathematics scores between the experimental and comparison classes (HPS)

The mean score differences were calculated by taking the mean score of the experimental class minus away the mean score of the comparison class. For instance, referring to Figure 1 above, for the Express stream classes, the difference in the mean score was 5.91 ($\approx$6) while the differences in the mean scores for F2004 and M2005 were 4.55 and 11.31 respectively. We also observed that in the Normal Academic stream classes, the differences in the mean scores consistently increased as students progressed from mid-year 2004 to final-year 2004 and to mid-year 2005.

Express stream. The data also revealed that there were significant differences in the mean scores between the experimental and comparison classes in the Express stream in both the M2004 (t [76] = 2.136, p = .036) and M2005 (t [77] = 3.105, p = .003). In addition, using the analysis of variance, it was also detected that the change in the mean scores from F2004 to M2005 for the experimental class is significantly different from the comparison class (F [1, 76] = 5.024, p = .028). In fact, from M2004 to F2004, both the experimental and comparison classes had decreased in their mean scores (but there was no significant difference was detected in the change between them), however, from F2004 to M2005, the experimental class had an increased in mean score while the comparison class's mean score had decreased (which resulted in a significant difference between these changes as shown in the above F value).

Normal Academic stream. The data revealed that there were no significant differences in the mean scores between the experimental and comparison classes for the students in the Normal Academic stream in all the three school-based mathematics examinations. In addition, the analysis of variance also did not detect any significant differences between the changes of the mean scores in the experimental class and the changes of the mean scores in the comparison class from either M2004 to F2004 or from F2004 to M2005. In fact, from M2004 to F2004, both the classes had decreased in their mean scores, and from F2004 to M2005, both classes had increased in their mean scores, however no significant change was detected between them.

Figure 9.1 also showed the number of intervention tasks done in each of the experimental classes during the 18 months period. Although it was hard to claim that doing communication tasks would bring about students performing better in their school-based mathematics examinations, we believed that the new tasks did not disadvantage students' mathematics learning at least as we could observe that students' results had either improved or significantly improved. We also believed that if the intervention period had been prolonged or extended, that was to say if teachers were to do more communication tasks, we might observe better results.

Non-high-performing school

At the beginning of the study, students' PSLE mathematics grades were used as an indication of the equivalent level of their achievement. The PSLE grades were analyzed using the chi-square test and the results showed that there was no significant difference in achievement level between the experimental and comparison classes in the Express stream (χ^2 [3, N = 43] = 3.370, p = .338) as well as in the Normal Academic stream (χ^2 [4, N = 46] = 5.592, p = .232).

The above results showed statistically that the experimental class and the corresponding comparison class (in the Express stream as well as in the Normal Academic stream) were of equivalent standard in mathematics achievement at the beginning of the study, as based on the PSLE mathematics grades obtained.

In a similar sense, the three sets of school's mathematics examination scores; M2004, F2004 and M2005 were also collected and the mean scores for the respective mathematics examinations for each class were evaluated and presented in Table 9.25.

Table 9.25. Mean scores of school-based examinations (NHPS)

Mean scores		M2004	F2004	M2005
Express	*Experimental*	57.71	50.81	61.14
	Comparison	63.95	57.95	58.59
Normal Academic	*Experimental*	66.00	70.08	55.04
	Comparison	55.09	44.91	36.86

The differences in the mean scores for each of the examination between the experimental and comparison classes were also calculated. Figure 2 showed the changes in the scores for both the Express and the Normal Academic classes. *Note and recall that: the difference in the mean score is calculated by taking the mean score of the comparison class subtracted away from the mean score of the experimental class.*

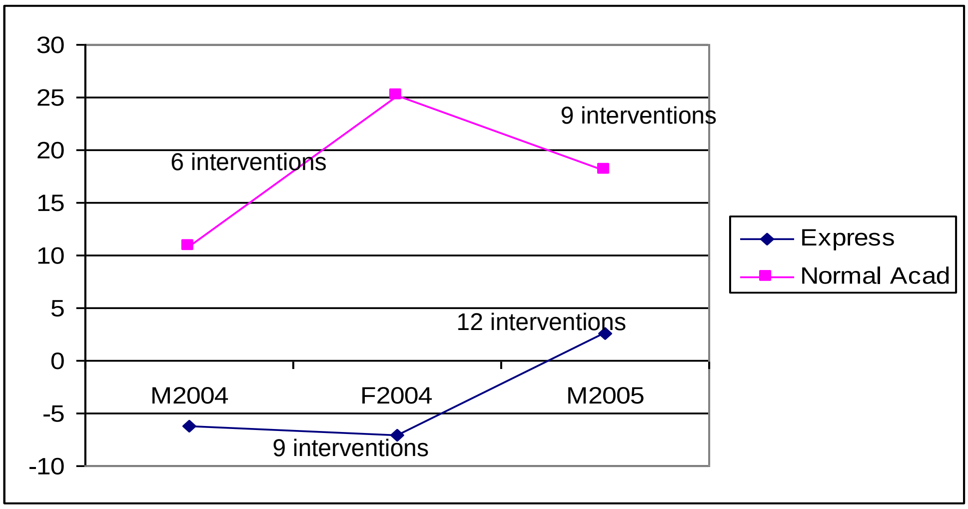


Figure 9.2. Differences in school-based exam mathematics scores between the experimental and comparison classes (NHPS)

Express stream. The data revealed that the by observing the difference in mean scores (which were negative values as shown in the above figure), the experimental class did worse than the comparison class in both M2004 and F2004 examinations. However, from F2004 to M2005, the experimental class did better than the comparison class. T-tests also revealed that there was no significant difference in the mean score between the experimental and comparison class in all the three examinations. Further analysis of variance also detected that the change in the mean scores from F2004 to M2005 for the experimental class was significantly different from the change that happened in the comparison class (F [1, 41] = 6.729, $p < .05$).

Normal Academic stream. The data in figure 2 revealed that the experimental class did better in all the three examinations compared to the corresponding comparison class. Although experimental class produced better results in all the three examinations compared to the comparison class, in M2005, the difference in mean score between them had shown slight decrease. T-tests revealed that there were significant differences in the mean scores between the experimental and comparison class in all the three examinations; M2004 (t [43] = 2.578, $p < .05$), F2004 (t [44] = 5.496, $p < .000$) and M2005 (t [44] = 4.324, $p < .000$). In fact, further analysis of variance also indicated that the change in the mean scores from M2004 to F2004, and from F2004 to M2005 for the experimental class were significantly different from the changes in the comparison class; the F values were respectively F [1, 43] = 37.393, $p < .000$ and F [1, 44] = 5.754, $p < .05$.

It was not difficult to observe from Figure 9.2 that the intervention seemed to have made a more positive impact on the Express students' results compared to students' in the Normal Academic stream. The total number of intervention tasks done in the Express experimental class was also observed to be more than the number of intervention tasks done in the Normal Academic experimental class. We believed that the number of intervention done was a crucial factor contributing the effect of the tasks on students' achievement. Moreover, the teacher in the Express class was much more serious and consistent in integrating the new tasks in her instruction compared to the teacher in the Normal Academic class.

Summary

From the start of implementing communication tasks starting around March 2004 until November 2004, most of the experimental classes only used journal writing tasks in mathematics classrooms except for the experimental Express class in the non-high performing school which started using oral presentation tasks after students' mid-year examination in May 2004. From January 2005 until the end of the intervention period in around May 2005, all the experimental classes used both journal writing and oral presentation tasks. The analysis of the students' school-based mathematics examinations did not show that experimental classes that significantly out-performed their corresponding comparison classes. In fact, both the experimental and the corresponding classes seemed to exhibit the same trend in their performance on the three school-based mathematics examinations results collected. However, all the experimental classes did show better scores in the M2005 examination compared to their corresponding comparison classes. The results appeared to indicate that in order for students to improve much more significantly, students need to engage in both written and oral communication tasks. Students' development in mathematics communication focusing in writing and speaking should come hand-in-hand and this seemed to might had made a difference in their performance in the conventional tasks.

9.4.3 Students' achievement in "journal writing task" tests

In this sub-study, it had adopted the use of journal writing as a form of written communication in mathematics classrooms. Unlike the conventional work that students did in the daily mathematics learning that usually dealt with routine computational skills, one of the purposes to use of communication tasks was to develop students' communication skills which included student's mathematical reasoning skills, and systematic and coherent presentational skills. In the study, journal writing intervention tasks were conceptually comprising of four aspects (see earlier discussion), however, the pre- and post-"journal writing task" tests were designed to only allowed students to demonstrate their mathematical reasoning of working on a solution or idea, and the coherence of their presentational skills. The pre- and post-"journal writing" tests were designed to be parallel so that students' changes (if any) in their abilities would be detected.

Both the pre- and post-"journal writing" tests consisted of two questions. The first question had two parts to it: 1(a) and 1(b). For each question 1(a), 1(b) and 2, it was graded using specific rubrics that looked into students' mathematical reasoning (MR) and presentation (Pres). The rubrics had five level of grading: Level 0, Level 1, Level 2, Level 3, and Level 4. Marks were awarded according to the level achieved by the students. Thus, the following will report results of the pre- and post-tests of the each question graded based MR and Pres and they were specifically in the pre-test: Pre1aMR, Pre1aPres, Pre1bMR, Pre1bPres, Pre2MR, Pre2Pres, and in the post-test: Post1aMR, Post1aPres, Post1bMR, Post1bPres, Post2MR and Post2Pres. For example, "Pre1aMR" would mean the results to the pre-test of question 1a testing on mathematical reasoning while "Post2Pres" would mean the results to the post-test of question 2 on presentation.

In this section, it will report first the results of the "journal writing" tests for the participating classes of the high-performing school followed by the participating classes from the non-high performing school.

High-performing school

Before the start of the intervention, students in the experimental and comparison classes took the same pre-test. Tables 9.26 and 9.27 showed the percentages of students achieving at each level for both the Express and Normal Academic classes, respectively. These tables allowed us to examine the equivalent level of the students between the experimental and comparison classes before the start of formal intervention in the experimental classes.

Table 9.26. Percentages of express stream students achieving at each level in rubrics grading for pre-test on "Journal Writing Task" (HPS)

Item/Class		Level 0	Level 1	Level 2	Level 3	Level 4
Pre1aMR	Experimental	0%	0%	0%	52.5%	47.5%
	Comparison	0%	0%	10.5%	55.3%	34.2%
Pre1aPres	Experimental	0%	45%	30%	22.5%	2.5%
	Comparison	0%	42.1%	36.8%	18.4%	2.6%
Pre1bMR	Experimental	0%	32.5%	37.5%	25%	5%
	Comparison	36.8%	44.7%	18.4%	0%	0%
Pre1bPres	Experimental	0%	10%	30%	50%	28.2%
	Comparison	7.9%	52.6%	34.2%	5.3%	0%
Pre2MR	Experimental	0%	100%	0%	0%	0%
	Comparison	7.9%	86.8%	5.3%	0%	0%
Pre2Pres	Experimental	0%	7.5%	92.5%	0%	0%
	Comparison	0%	10.5%	81.6%	7.9%	0%

Referring to Tables 9.26, first let us examine by comparing the results of the experimental and the corresponding comparison class in the Express stream. The data revealed that for both streams, the experimental and the corresponding comparison class were quite equivalent in their level of achievement in each aspect of the measurements (e.g. Pre1aMR, Pre1bPres, etc). For example in Table 23 under the aspect Pre1aMR, the total percentages of students' scores were around level 3 and 4 for both the experimental and comparison classes. Similarly, under the aspect Pre1aPres, the total percentages of students scoring Level 1 and 2 were 75% for the experimental class and 78.9% for the comparison class. This phenomenal of equivalence between the experimental and comparison classes could be observed for aspect in Pre2MR and Pre2Pres except for Pre1bMR and Pre1bPres. In fact, Mann-Whitney U tests further testified that there were significant differences in students' performance in these 2 aspects between the experimental and comparison classes, where in Pre1bMR yields U [40, 38] = 254.00, $p <$ .000 and in Pre1bPres yields U [40, 38] = 222.00, $p < .000$.

Table 9.27. Percentages of normal academic stream students achieving at each Level in rubrics grading for pre-test on "Journal Writing Task" (HPS)

Item/Class		Level 0	Level 1	Level 2	Level 3	Level 4
Pre1aMR	Experimental	2.5%	5%	25%	47.5%	20%
	Comparison	7.5%	5%	17.5%	47.5%	22.5%
Pre1aPres	Experimental	0%	50%	35%	15%	0%
	Comparison	7.5%	60%	22.5%	10%	0%
Pre1bMR	Experimental	42.5%	42.5%	10%	5%	0%
	Comparison	50%	27.5%	20%	2.5%	0%
Pre1bPres	Experimental	2.5%	67.5%	25%	5%	0%
	Comparison	10%	62.5%	15%	12.5%	0%
Pre2MR	Experimental	25%	75%	0%	0%	0%
	Comparison	7.5%	92.5%	0%	0%	0%
Pre2Pres	Experimental	2.5%	30%	67.5%	0%	0%
	Comparison	0%	17.5%	82.5%	0%	0%

Likewise, from Table 9.27, first examining by comparing the results of the experimental and the corresponding comparison class in the Normal Academic stream, we can observe that the scores between them were quite equivalent except for aspect Pre2MR where most students in the scored Level 0 and Level 1 in the experimental class whereas in the comparison class, almost all the students scored Level 1. Further analysis revealed that there was significant difference between students in the experimental and comparison class for aspect Pre2MR with U [40, 40] = 660.00, $p = .035$.

After about 18 months of implementing communication tasks in the experimental classes, a post-test was again administrated to all the four classes. This study will report results by comparing the pre- and post-scores obtained by the students within the class as well as comparing the post-scores of students between the experimental and comparison classes. The pre- and post-results in Table 9.28 revealed that in the experimental Express class, students had significantly improved in their performance for the aspect in mathematical reasoning in question 2 ($Z = -5.610$, $p < .000$), and presentation skills in question 1b ($Z = -4.912$, $p < .000$) and 2 ($Z = -4.983$, $p < .000$) but significantly decreased in scores for mathematical reasoning in question 1a ($Z = -3.207$, $p < .005$). The Z values here were based on using the Wilcoxon Signed Rank test.

Table 9.28. Percentages of express stream students achieving at each level in rubrics grading for pre- and post-test on "Journal Writing Task" (HPS)

Item/Test		Level 0	Level 1	Level 2	Level 3	Level 4
1aMR	Post1aMR	2.5%	2.5%	37.5	25%	32.5%
	Pre1aMR	0%	0%	52.5%	47.5%	47.5%
1aPres	Post1aPres	2.5%	47.5%	30%	17.5%	2.5%
	Pre1aPres	0%	45%	30%	22.5%	2.5%
1bMR	Post1bMR	5%	5%	72.5%	17.5%	0%
	Pre1bMR	0%	32.5%	37.5%	25%	5%
1bPres	Post1bPres	0%	0%	25%	75%	0%
	Pre1bPres	0%	10%	30%	50%	28.2%
2MR	Post2MR	0%	0%	30%	47.5%	22.5%
	Pre2MR	0%	100%	0%	0%	0%
2Pres	Post2Pres	0%	2.5%	25%	50%	22.5%
	Pre2Pres	0%	7.5%	92.5%	0%	0%

As for the experimental Normal Academic class (see Table 9.29), students had significantly dropped in their performance for aspect concerning mathematical reasoning in question 1a (Z = -5.058, *p* < .000), and presentation in question 1a (Z = -4.138, *p* < .000) and 2 (Z = -4.059, *p* < .000) but improved in presentation in question 1b (Z = -4.307, *p* < .000). This was no surprise because this experimental Normal Academic class was not very serious about their learning all the time in class and classroom observation had also found that the teacher tried very hard to implement the intervention tasks to the students throughout the whole intervention period. As for the comparison Express class, there were no consistent results because the Signed Rank test revealed that the class had significantly improved in mathematical reasoning in question 2 (Z = -3.809, *p* < .000) but significantly dropped question 1a (Z = -3.903, *p* < .000), and significantly improved in presentation in question 1b (Z = -4.804, *p* < .000) but significantly dropped in question 1a (Z = -3.069, *p* < .005). As for the comparison Normal Academic class, students had significantly performance worse in the post-test for aspect in mathematical reasoning in question 1a (Z = -4.219, *p* < .000) and 1b (Z = -2.841, *p* < .005), and presentation in 1a (Z = -3.749, *p* < .000).

Table 9.29. Percentages of normal academic stream students achieving at each level in rubrics grading for pre- and post-test on "Journal Writing Task" (HPS)

Item/Test		Level 0	Level 1	Level 2	Level 3	Level 4
1aMR	Post1aMR	25%	47.2%	27.8%	0%	0%
	Pre1aMR	2.5%	5%	25%	47.5%	20%
1aPres	Post1aPres	25%	72.2%	2.8%	0%	0%
	Pre1aPres	0%	50%	35%	15%	0%
1bMR	Post1bMR	55.6%	8.3%	36.1%	0%	0%
	Pre1bMR	42.5%	42.5%	10%	5%	0%
1bPres	Post1bPres	19.4%	0%	2.8%	41.7%	36.1%
	Pre1bPres	2.5%	67.5%	25%	5%	0%
2MR	Post2MR	47.2%	25%	25%	2.8%	0%
	Pre2MR	25%	75%	0%	0%	0%
2Pres	Post2Pres	47.2%	30.6%	19.4%	2.8%	0%
	Pre2Pres	2.5%	30%	67.5%	0%	0%

In addition, by comparing the post-results between the experimental and the corresponding comparison classes, the data revealed that in the Express stream, the experimental class did significantly better in all the components: Post1aMR (*U* [40, 37] = 501.00, *p* < .05), Post1aPres (*U* [40, 37] = 485.00, *p* < .005), Post1bMR (*U* [40, 3] = 257.00, *p* < .000), Post1bPres (*U* [40, 37] = 330.00, *p* < .000), Post2MR (*U* [40, 37] =

374.00, $p < .000$), and Post2Pres (U [40, 37] = 349.00, $p < .000$) compared to the comparison class, whereas in the Normal Academic stream, the experimental class did significantly better in only 2 components compared to the comparison class and they were Post1bMR (U [36, 28] = 360.500, $p < .05$) and Post1aPres (U [36, 28] = 353.500, $p < .05$). This is not unexpected because, naturally the experimental class had more opportunities to be engaged in mathematical reasoning and presentation when they were working on the journal writing tasks.

Non-high-performing school
In a similar way to the high-performing school, before the start of the intervention, students from the non-high performing school both in the experimental and comparison classes took the same pre-test. The following Table 9.30 and 9.31 showed the percentages of students achieving at each level according to the grading rubrics for both Express and Normal Academic classes respectively. These tables allowed us to examine the equivalent level of the students' ability between the experimental and comparison classes before the start of the formal intervention in the experimental classes.

Table 9.30. Percentages of express stream students achieving at each level in rubrics grading for pre-test on "Journal Writing Task" (NHPS)

Item/Class		Level 0	Level 1	Level 2	Level 3	Level 4
Pre1aMR	Experimental	0%	4.8%	14.3%	47.3%	33.3%
	Comparison	4.5%	9.1%	13.6%	40.9%	31.8%
Pre1aPres	Experimental	0%	57.1%	23.8%	14.3%	4.8%
	Comparison	0%	23.8%	23.8%	33.3%	19%
Pre1bMR	Experimental	9.5%	57.1%	19%	14.3%	0%
	Comparison	9.1%	50%	40.9%	0%	0%
Pre1bPres	Experimental	0%	47.6%	38.1%	14.3%	0%
	Comparison	4.5%	27.3%	54.5%	13.6%	0%
Pre2MR	Experimental	23.8%	76.2%	0%	0%	0%
	Comparison	9.1%	90.9%	0%	0%	0%
Pre2Pres	Experimental	0%	38.1%	61.9%	0%	0%
	Comparison	0%	19%	81%	0%	0%

Data in Table 9.30 revealed that in the Express stream, the experimental class is quite equivalent to the corresponding comparison class in all the aspects except for Pre1aPres which had a significant difference in performance between these two classes based on Mann-Whitney U test (U [21, 21] = 127.00, $p < .05$). To explain further about the equivalence, observe that in aspect Pre1aMR the total percentages of students in the experimental class reaching level 3 and 4 was 80.6% compared to 72.7% in the comparison class; for aspect Pre1bPres the total percentages of students in the experimental class reaching level 1 and 2 was 85.7% which is quite equivalent to 81.8% in the comparison class; as for aspect Pre2MR and Pre2Pres, all students in the experimental and comparison classes reached 100% in achievement in level 0 and 1, and level 1 and 2 respectively.

Likewise, in Table 9.31, the data revealed that in the Normal Academic stream, the experimental class is also quite equivalent to the corresponding comparison class based on the same analysis as the above. However, equivalence did not show for aspect Pre1bMR (U [23, 21] = 145.500, $p < .005$) and Pre2Pres (U [21, 21] = 147.00, $p < .05$).

Table 9.31. Percentages of normal academic stream students achieving at each level in rubrics grading for pre- and post-test on "Journal Writing Task" (HPS)

Item/Class		Level 0	Level 1	Level 2	Level 3	Level 4
Pre1aMR	Experimental	34.8%	21.7%	30.4%	13%	0%
	Comparison	57.1%	28.6%	9.5%	4.8%	0%
Pre1aPres	Experimental	19%	52.4%	23.8%	4.8%	0%
	Comparison	0%	57.1%	38.1%	0%	4.8%
Pre1bMR	Experimental	8.7%	73.9%	17.4%	0%	0%
	Comparison	42.9%	52.4%	4.8%	0%	0%
Pre1bPres	Experimental	0%	69.6%	30.4%	0%	0%
	Comparison	4.8%	81%	14.3%	0%	0%
Pre2MR	Experimental	30.4%	69.6%	0%	0%	0%
	Comparison	14.3%	85.7%	0%	0%	0%
Pre2Pres	Experimental	0%	28.6%	71.4%	0%	0%
	Comparison	0%	61.9%	38.1%	0%	0%

A post-test was administrated at the end of 18 months of intervention to all the experimental and comparison classes. First, we will examine the pre- and post-results of each class and then compare the results of the post-tests between the experimental and the corresponding comparison classes. The pre- and post-results data in Table 9.32 revealed that the experimental Express class students had significantly improved in their performance for aspect measuring students' mathematical reasoning in question 2 ($Z = -3.666$, $p < .000$), students' presentation skills in question 1b ($Z = -3.681$, $p < .000$) and 2 ($Z = -2.815$, $p < .005$) but dropped in score for mathematical reasoning in question 1a ($Z = -2.450$, $p <.05$).

Table 9.32. Percentages of express stream students achieving at each level in rubrics grading for pre- and post-test on "Journal Writing Task" (NHPS)

Item/Test		Level 0	Level 1	Level 2	Level 3	Level 4
1aMR	Post1aMR	0%	5%	55%	35%	5%
	Pre1aMR	0%	4.8%	14.3%	47.3%	33.3%
1aPres	Post1aPres	5%	50%	35%	10%	0%
	Pre1aPres	0%	57.1%	23.8%	14.3%	4.8%
1bMR	Post1bMR	10%	40%	30%	15%	5%
	Pre1bMR	9.5%	57.1%	19%	14.3%	0%
1bPres	Post1bPres	0%	0%	0%	65%	35%
	Pre1bPres	0%	47.6%	38.1%	14.3%	0%
2MR	Post2MR	0%	15%	25%	45%	15%
	Pre2MR	23.8%	76.2%	0%	0%	0%
2Pres	Post2Pres	0%	10%	40%	45%	5%
	Pre2Pres	0%	38.1%	61.9%	0%	0%

As for the experimental Normal Academic students (see Table 3.33), the data revealed that students had significantly improved in their scores for aspect measuring mathematical reasoning in question 1a ($Z = -2.930$, $p < .005$) and 2 ($Z = -2.714$, $p < .005$) but decreased in question 1b ($Z = -2.138$, $p < .05$); and significantly improved in scores for presentation skills in question 1b ($Z = -3.138$, $p < .005$) but decreased in question 2 ($Z = -3.464$, $p < .005$). Both the comparison classes in the Express stream and in the Normal Academic stream had inconsistent results. To illustrate further first consider the comparison Express class. Comparing the pre- and post-results, the class had significantly improved in mathematical reasoning in question 2 ($Z = -3.8041$, $p < .000$) but dropped in scores in question 1a ($Z = -3.914$, $p < .000$); and significantly improved in presentation skills in question 1b ($Z = -3.214$, $p < .005$) but dropped in scores in question 1a ($Z = -3.210$, $p < .005$). This phenomenon was also observed in the comparison

Normal Academic class. By comparing the pre- and post-results, this class had also significantly improved in mathematical reasoning in question 2 (Z = -2.121, $p < .05$) but decreased in question 1b (Z = -2.121, $p < .05$); and significantly improved in presentation skill in question 1b (Z = -3.727, $p < .000$) but decreased in question 1a (Z = -3.491, $p < .000$).

Table 9.33. Percentages of normal academic stream students achieving at each level in rubrics grading for pre- and post-test on "Journal Writing Task" (NHPS)

Item/Test		Level 0	Level 1	Level 2	Level 3	Level 4
1aMR	Post1aMR	0%	20.8%	29.2%	50%	0%
	Pre1aMR	34.8%	21.7%	30.4%	13%	0%
1aPres	Post1aPres	0%	91.3%	8.7%	0%	0%
	Pre1aPres	19%	52.4%	23.8%	4.8%	0%
1bMR	Post1bMR	41.7%	41.7%	16.7%	0%	0%
	Pre1bMR	8.7%	73.9%	17.4%	0%	0%
1bPres	Post1bPres	20.8%	0%	4.2%	41.7%	33.3%
	Pre1bPres	0%	69.6%	30.4%	0%	0%
2MR	Post2MR	12.5%	70.8%	16.7%	0%	0%
	Pre2MR	30.4%	69.6%	0%	0%	0%
2Pres	Post2Pres	13%	73.9%	13%	0%	0%
	Pre2Pres	0%	28.6%	71.4%	0%	0%

In addition, by comparing the post-results between the experimental and the corresponding comparison classes, the data revealed that in the Express stream, the experimental class did significantly better than the comparison class in two aspects namely Post1aMR (U [20, 22] = 57.00, $p < .000$) and Post1bPres (U [20, 22] = 147.00, $p < .05$), whereas in the Normal Academic stream, the experimental class also did significantly better than the comparison class in Post1aMR (U [24, 20] = 41.00, $p < .000$), Post1bMR (U [24, 20] = 160.00, $p < .05$) and Post1aPres (U [23, 20] = 136.50, $p < .005$).

Summary

The pre- and post-"journal writing" tests were intended to measure the students' ability to demonstrate their mathematical reasoning skills, and their systematic and coherent presentational skills which were considered in this study as two important aspects in the development of students' communication proficiency. The results revealed that students in the experimental classes whom had been receiving journal writing for about 18 months showed that they were generally more competent in mathematical reasoning and presentational skills compared to their corresponding comparison classes in the post-test.

9.4.4 Interviews with teachers and students

Interviews were conducted for all the participating teachers in the experimental classes and selected students from all the participating experimental classes at the end of the intervention period in May 2005. A total of five teachers and 30 students (6 from each participating class) were interviewed. The students were selected by their respective mathematics teachers. In order to get a more generalized results from the interviews, students selected comprised of high ability, average ability and low ability caliber. All the teachers were interviewed individually whereas students from the same class were interviewed three (1 from high ability, 1 from low ability and 1 low ability) at a time. Each interview session for teachers lasted between 25 to 58 minutes, while student interview took about 11-28 minutes in each session. All the interviews were audio recorded and transcribed thereafter.

One of the main reasons of the interviews was to get direct information and confirmation from both the teachers and students about issues pertaining to the use of communication tasks in the teaching and learning mathematics. The interview questions for both teachers and students were categorized into three main parts: (a) Own experience and understanding about journal writing and oral presentation tasks, (b) Opinions and comments about students' ability to perform journal writing and oral presentation tasks, and (c) Comments and suggestions for effective and efficient implementation of journal writing and oral presentation tasks during teaching.

In this section, it will first report the results from the interviews with teachers followed by the results from the interviews with students.

About experience and understanding about the new tasks
Participating teachers and students were relatively new to the exposure of communication tasks in this study. After about 18 months of doing the communication tasks in the daily mathematics classrooms, the study would like to find out about teachers' and students' general attitudes and beliefs toward the use of these new tasks as well as to share their knowledge and perceptive about these new tasks. The questions asked were to find out whether both teachers and students (i) understood the differences between these new tasks and the conventional ones that students were currently working on in their daily learning, (ii) believed these new tasks were able to benefit students' learning or bring about more harm to students' learning, (iii) had any difficulties in using and doing these new tasks, and (iv) thought that these new tasks could actually be feasibly used and how often could be used in the daily mathematics classrooms.

1. Understanding about the new tasks
Views from the teachers. Generally all the teachers believed that the communication tasks were non-routine and required students to explicitly explain, describe, and illustrate their understanding of mathematics contents using sustained writing and speech. Based on the teachers' own experiences and understanding, a list of differences between the traditional or conventional tasks and the new tasks that were quoted by the teachers were as followed:

Conventional tasks	Communication tasks
• Questions usually involve figures, numbers and often deal with problem-solving and calculation. • Questions are often routine, rote, procedural and mechanical. • The purpose is more for drill and practice. • Questions are structured to achieve the 'O' level syllabus. • Emphasize more on the outcome/ results of the questions/problems. • Questions require students to merely write down the mathematical steps and then answers without much explanation.	• Questions allow more in-depth thinking. • Questions give students opportunity to explain how they derive the solutions of the problems or why they use certain methods in their solutions. • Get students to reflect what they have learnt. • Get students to think about concepts taught. • Ask students to explain in proper English statements in both writing and oral. • Allow students to think about their mathematical language when they are writing and speaking.

Views from the students. Similarly, students were also asked to about their own understanding on the new tasks. All the students interviewed generally had clear understanding of the differences between the new tasks and the conventional/normal work that they did in class. Students explained that the journal writing task was more than just work on the solutions of questions; it required them to elaborate and explain why and how they worked on their mathematics questions.

In addition, students expressed that journal writing also allowed them to pen down their feelings, their likes and dislikes, and their frustration to their teachers. Similarly students also understood that oral task required them to speak out and share with their classmates and teachers their reasoning to mathematical workings and solutions.

The followings were students' perceptions about the features of the new tasks, and differences between the new tasks and the conventional tasks that they had worked on based on their own experiences and understanding:

- *This one must think more.*
- *We can write our problems or anything we don't know here.*
- *This one test our understanding more.*
- *This one more like speaking up your mind.*
- *Have to explain more.*
- *Explain more in detail, writing helps us to recall more.*
- *Need to elaborate more.*
- *To crack our brain to think about it.*
- *A lot of writing and explanation.*
- *They are more like thinking questions.*
- *Oral you get immediate response.*
- *Doing these new things is more than getting the results but process of getting it.*
- *These new tasks ask what the meaning is; we can tell what we really think about… For the normal ones, they ask only like math questions, solve it, solve for x.*
- *These ones you need to explain more clearly unlike the normal ones you write steps only and not in words, these ones are like explain more clearly.*
- *They ask we all to explain the meaning and make us more understand about the topic.*
- *You can understand better by explaining to other people and we can know our mistakes, the common mistakes we make.*
- *Traditional task only do those working and we never explain. The new tasks let us explain more in details in the oral presentation and we can also write out in the journal writing.*
- *Writing down our feelings is not so embarrassing.*
- *We have to reason our working.*

2. Beliefs about the benefits (if any) by doing the new tasks

Views from the teachers. Both the teachers teaching the Normal Academic classes felt that the new tasks did not seem to have benefited some of their students directly because of the general 'bad' attitudes of the students. The teachers felt that it was not because of the nature of the new tasks that did not benefit their students but because these students were already not very serious about their learning whether it was about using traditional task or working on the new tasks. One of the teacher commented that the new tasks would benefit only some of her students because some of her students were not serious in their work and they were also rather temperamental:

- *…some of the students are temperamental*
- *…if they really want to write, they can but if they don't want to, but then….*

The above teachers believed that it was the attitudes of her and his students that generally needed to be improved in order for them to learn better.

For the other two teachers teaching the Express streams felt that the new tasks would benefit their students because the new tasks gave students the opportunity to engage in learning mathematics in-depth and understanding concepts better. Some of the comments included:

- *…math is not just getting the answer correct but how to get the answer*
- *A lot of things we do never question, just follow the rules, this is not enough*

Besides the teachers in the Normal Academic stream commenting that it was the attitudes of her and his students that needed to be improved in order for them to learn better, generally teachers did recognized and believed that the new tasks would benefit students' mathematics learning.

In addition, due to the nature of communication tasks, which had to do with writing and speech, and therefore it was not surprising that all the teachers felt that the disadvantage of using the new tasks was that if students' command of the English language was weak, students might have problems in writing and speaking. However teachers also believed that practicing good writing and oral presentation should be made as part of the process of mathematics learning because learning of mathematics should not merely focused in "solve and answer." The general consensus was that although students with weak command of language might be disadvantaged at the beginning, with proper guidance and practice, they should overcome this problem in the long run.

In addition, teachers also believed that if students were weak in communication due to weak in language, all the more students needed to develop and improve in this aspect even in the mathematics classrooms. The advantages of doing communication tasks were often mentioned in the perspective of students. For examples, teachers made the following statements:

- *Mathematical language is emphasized, so get students to think about the language.*
- *I appreciate writing because students get to express themselves; writing will elicit some of their hidden responses.*
- *Give everyone a fair chance to tell you what they think about mathematics and their personal feeling.*
- *Help them to arrange their ideas like for example they know the steps but do not know why they get it, thus facilitate their learning, give them a clearer picture of what they are doing and not just doing things blindly.*
- *Writing is a very comfortable environment, just face paper.*
- *Oral task develops students holistically because you must think through on what you want to say and then how you say it; it is a higher learning skill compared to writing.*
- *Allows students to reflect more on what they are learning.*
- *I can use these tasks as filler in my mathematics lessons.*
- *Able to identify who are the weaker students and to go through students' problems that they encountered.*
- *Students' writings, that's where you know where and what kind of problem they are facing.*

Views from the students. All the students interviewed believed that the new tasks were helpful in their learning of mathematics although some felt that doing these new tasks could be difficult at the beginning.

The followings are students' explanation on how the new tasks had and would help them in their mathematics learning, based on their own experiences:

- *When I write I will remember more things…when I talk, more things can be taken down and I can remember more of my weaknesses.*
- *When we explain, we get to know the method more.*
- *Explain more in detail, writing helps us to recall more.*
- *Journal writing allows us to write the things we don't understand down but when we do worksheet questions when we don't understand or maybe we just try the*

questions and if the teacher didn't explain further then you will not really understand. But if you write down then you will understand more about your weaknesses.

- *In this way we can know how to think about it, how to get the answer, is not important in the results but important how to get it.*
- *During oral presentation, I can express myself in a more detailed manner.*
- *Can help us think more about what we have learned and then apply it on our work.*
- *You can understand better by explaining to other people and we can know our mistakes.*
- *We can see more clearly in our workings and we can learn from our mistakes.*
- *Last time when we don't know we just skip or copy. But now if you don't know your work, you have to find out because you have to write down. As you write down, you will find out your own mistakes and what you are not sure.*
- *Must understand first before you can write anything or say anything.*
- *It makes us understand better… it makes you clearer.*
- *We can learn more different methods to do problems.*
- *By writing down our reflection we can improve our weak point.*
- *Writing you can understand better and know what we understand, we know where which level we are in.*

3. Difficulties encountered

Views from the teachers. The difficulties and challenges that the teachers faced were often the time factor, i.e. teachers felt that they did not have time to implement the tasks in class. This was quite understandable because at the beginning of the intervention, most teachers felt that the new tasks were just additional work for both their teaching and students' learning. That was to say, the new tasks were seen as add-on to the regular tasks that they had to do which were already pre-planned in their scheme of work earlier in the year. The teachers therefore felt that it was a great challenge to them to be able to integrate these new tasks into their already pre-planned scheme of work.

The following were some of their responses to the difficulties and challenges they faced when administrating the new tasks in their own classrooms:

- *Time constraint is a factor.*
- *No time to do, time constraint.*
- *Having time to do and read what they read.*
- *To give them regular feedback.*
- *The most difficult is really giving them feedback on time.*
- *To spend time to actually sit down and mark and write as a response to what they write; this is a challenge.*
- *Journal writing is something sometimes open, so basically it's new and not clear how to encourage them to write more (to substantiate their writing) and give them so-called a model answer is difficult.*

Unlike in the beginning when the teachers first encountered communication tasks when these new tasks were relatively new to them, after about three school terms, all the teachers interviewed were comfortable in using the new tasks in their daily teaching now; they did not express any extreme or apparent problem and concern. During the intervention period, the teachers were often advised by the researcher that before using the new tasks in the classroom, always asked yourself what was the purpose of using the new task, what were that I wanted to find out about my students' understanding, could the prompts in the new tasks be able to let me find out about what I wanted to find out? The teachers were made to understand that integrating the new tasks in the daily teaching was the same as integrating the conventional work for the students during daily teaching. The two important questions were: Why and for what purpose do I want to give this piece of work to my students? And what learning experience outcome do I want my students to have by doing this piece of work? One of the teachers stated that:

- *After sometimes, you actually get quite familiar….knows what to do.*

Another teacher also added that:

- *Later on, I am clear what to look for, what to ask.*

Teachers were also asked about the difficulties that their students faced in the process of doing the new tasks in the mathematics classroom.

Views from the students. The study also intended to find out what were the difficulties students had when they worked on these new tasks for the first time and whether their subsequent experiences helped them to overcome the difficulties they first encountered.

All the students except for two became more confident after working on the tasks for about four months (for some students) or 16 months (for some students) compared to the first time they encountered these tasks. However all the students believed that with more practices and gaining more experiences would help them improve their skills in writing and speech. The difficulties faced by the students were as follows:

- *Don't know what to write…because we must write in details… in sentences.*
- *Don't know how to start.*
- *Don't know how to explain the steps out.*
- *Sometime we don't know how to put it in words.*
- *Sometime quite difficult to express everything.*
- *Fear to talk in front of the whole class.*
- *You get scared that maybe if you write or say wrongly, then people might be affected. You think that this is correct and they might also take the wrong idea.*
- *My English is not good so I can't write out what actually I think and very difficult to express myself the work through words having to write out.*
- *Oral presentation sometime you get scared, nervous. Then the writing is, sometime you want to write down but you don't know what how to put it.*

4. Feasibilities of using new tasks in mathematics classrooms

Views from the teachers. All the teachers interviewed except for only one teacher teaching the Normal Academic class felt that the new tasks cannot be substituted with the traditional and conventional tasks because of the examinations that the students had to take. This teacher felt very strongly that 'drill and practice' was still the most important part in mathematics learning for preparing his students for the examination.

The teacher felt that unless students' communication skills were to be formally assessed or reflected in the assessment system, then communication tasks should be incorporated into the classrooms. Or else there was no good reason why conventional tasks should be replaced with the new tasks. Since students needed to produce good results only in the examination where the examination highly emphasized mainly in computational skills, then one should teach accordingly so as to enable students to get good results. The reason was that for the sake of producing good results in the examinations; *"no matter what, students take the exam and they need to produce results!"*

As for the rest of the other four teachers, they generally felt that as long as the new tasks were appropriately designed to measure students' specific learning outcome, then it should be fine. In addition, they also felt that the some of the traditional or conventional tasks could be replaced by the new tasks provided these new tasks were to also become part of the assessment modes for the students. These teachers made the following comments:

- *Because the mode is still in your own words, so assuming that they write all they know, then I think it will be reliable. I think it will be good for some of these traditional exercises be replaced with these new tasks. Probably about 10% of the topics done.*
- *Actually the sec ones right now have an assessment on one of the journal writing, I guess it is not a big problem. Say about two per term that will be quite manageable. New tasks measure different aspects of their learning, learning the understanding of the concepts as in how to explain in words.*
- *A reliable source to find out from them whether students internalized what they have learnt rather than just solve. I think should be less than 10% of those regular exercises be replaced with these new tasks.*

In addition, all the teachers felt that if the new tasks were to be integrated into the daily mathematics curriculum, then the new tasks would become part of students' work and therefore it did not matter how often students work on these new tasks but for what purpose students need to work on these new tasks. However at the beginning when everybody was new to the new tasks, for example, journal writing, it could be done at least once in either two weeks or three weeks. While as for formal oral presentation, two teachers expressed that it could be done once a month, one teacher felt that it could be done once a term and the other teacher believed that oral was something that could be done everyday.

Most of the teachers believed that the new tasks could be implemented anytime in the curriculum; before, middle or end of a topic. One teacher expressed that: *"…. after every topic or after every two topics, sometimes it is within the topic itself."*

Views from the students. All the students said 'yes' they would think that the new tasks could be used in mathematics teaching and would support more of these new tasks to be done in mathematics classrooms except for two students who expressed that however it might be difficult for students to change or adapt themselves to do these tasks. For example one of them explained:

- *Is from like a very long time we do practices and all those things and then we need to do these suddenly is so hard lah! If I think for new people like in primary one like that you start doing these until now they will adapt.*

However, when the interviewer asked a further question about what if teachers were to guide and train them slowly in doing these new tasks, both the above two students agreed that they would support in this case.

Here were some of the comments made by the students regarding if the use of the new tasks were feasible and if they would the use of these tasks in their learning:

- *Support because may be by doing this kind of new task you can improve more your mathematics, it's easier for them to understand.*
- *Teacher may know the student better.*
- *More interesting.*
- *So the teacher knows if we are doing the right thing or wrong thing.*
- *At least they don't understand also after doing these then they will understand even better.*
- *Because when we do our working and the question here says we need to explain why, so if we understand really well we will know how to explain.*
- *It gives us plus the know what's the weak point.*

- *We can think like the teacher.*
- *Half yes half no. Yes because you can understand more, understand [more] clearer, faster, easier. No because must write a lot.*
- *Help us train our confidence.*
- *Can interact with the teacher more.*

In addition, some students also gave their comments on how often these new tasks could be used in the mathematics classroom. For those who commented, offered some suggestion to how frequent their teachers could integrate the new tasks during classroom teaching and students' opinions ranged from 'once a month' to 'everyday'.

The following tables (9.34 and 9.35) listed students' proposed frequency for the use of journal writing and oral task respectively in mathematics classrooms:

Table 9.34. Proposed frequency for use of journal writing in Mathematics classrooms

Number of students responded	Journal Writing Frequency
5	Once a week
5	Twice a week
3	Once every 2 weeks
2	Once a month
2	Twice a month
3	2 or 3 times a month
2	Once every topic
2	After one chapter
1	Everyday

Table 9.35. Proposed frequency for use of oral task in Mathematics classrooms

Number of students responded	Oral Task Frequency
9	Once a week
4	Once every 2 weeks
1	Once a month
2	Twice a month
2	After one chapter
1	As often as possible
1	Everyday

Opinions and comments about students' ability to work on the new tasks
This category of questions were to find out from teachers (i) how they felt about their students' ability to performance on the new tasks, (ii) what kind of difficulties their students encountered and, how we could help students overcome their difficulties to make them learn better in using the new tasks. In this section, it will also report comments from the students' perspective about their own abilities.

1. Students' ability to do new tasks

Views from the teachers. Interestingly, both the teachers in the non-high performing school estimated a lower percentage of students' ability to engage in these new tasks compared to the teachers in the high-performing schools. It seemed like the teachers in the high-performing schools believed that their students were more capable than those from the non-high performing school. Each individual teacher gave their own estimation, the percentages of their students being able to do journal writing task and oral presentation task. Table 9.36 provided the percentages given by the individual teachers based on their own view of their own students' ability.

Table 9.36. Percentages given by individual teachers based on their own view of students' ability

Schools and streams		Ability to do journal writing	Ability to do oral presentation
High-performing school	Express	80%	40%
	Normal academic	75%	20%
Non-high-performing school	Express	40%	10%
	Normal academic	20%	10%

According to the teachers, students were generally interested in the new tasks because the new tasks were non-routine. One teacher commented: "*they are not like solve questions from 1a to 1f kind of thing."* The teachers believed that probably there were two main reasons for students who detested in doing the new tasks. One reason had to do with students' attitude. One teacher expressed that: *"…whether it is journal writing or normal day work, they just don't want to do it!"* The other reason was likely to have to do with students having difficulty with the language, quote: *"…language is probably a hinder for them to write or speak."*

Due to the fact that the new tasks were unfamiliar and required different learning ability from the students, the teachers were not surprised about their students' initial 'dislike' about the new tasks and the types of difficulties encountered by them. They believed that students' ability in their command of the English language was a crucial factor in determining their ability to perform well in the new tasks. Based on their daily classroom observations, the teachers had identified the difficulties encountered by their students and they were as follow:

- *Language problem; they can write down or use mathematical notation but not explain in English form.*
- *They can't express themselves well, from the thoughts to the words portion. They do not know how to express what they don't understand. They are contorted, so those are not proficient in using math language.*
- *Secondary math is often just working and the answer; so asking them to elaborate is new.*
- *Students are not used to verbalize math concepts in words.*
- *Students are shy and reserved; they are terrified to have to speak to the whole class.*
- *Sometimes they are stuck, they don't know what you expect them to write or say.*
- *Students are afraid of making mistakes.*

Views from the students. The comments given by the students were rather consistent with those remarks told by the teachers that one of the major problems they had in working on the new tasks was due to their weak command of the English language. Students' comments on their difficulties working on the new tasks were reflected in the earlier above paragraphs on page 37.

Nevertheless, students also expressed that after one year and four months, they had become more confident in working on the new tasks. In fact, **a**ll the students interviewed said yes and agreed that they were more comfortable compared to the first time when they worked on these tasks. They claimed that they now understand what was expected, knew how to start or begin writing and speaking after some practices. One common comment was: *more comfortable now after a year because we have experience already.*

Students were also asked about their ability to complete the new tasks on time. Based on the responses given by the students, the time taken to complete a journal writing task is summarized in Table 9.37.

Table 9.37. Responses of students on time taken to complete a journal writing task

Shortest time taken to complete the task	5 to 15 minutes
Longest time taken to complete the task	30 minutes to 1 hour or (take home for 1 day to complete)
Average time taken to complete the task	20 to 25 minutes

Generally, they did not have much problem in completing the new tasks on time whether the task was to be done in class or as homework. They commented that they were able to submit their work promptly as instructed by the teachers. Consistently teachers also commented that collecting students work at a pointed time in the classroom was never the problem however, teachers further remarked that they did not like the idea of students having to bring work back. The reason was that students often gave excuses of not being able to bring work back to school.

2. Facilitate students working on new tasks

Views from the teachers. The teachers have listed out a considerable number of significant points and actions which they had taken during the intervention period and believed would help students improve their performance in the new tasks and eventually in their learning of mathematics. The teachers' descriptions included managing students in class, explaining clearly expectations to students, and providing appropriate teaching instructions.

The following were some of the quotes extracted from the teachers' interview sessions:

- *First of all, teacher has to speak the right language.*
- *Give them a 'pleasant environment', a non-threatening environment with peers' support, an environment that they can speak freely.*
- *Give students a positive and supportive environment. Form rules, first thing is respect.*
- *They need to know that it's ok to make mistakes; everyone is learning.*
- *You've got to encourage them, you've got to convince them, show them in a really feasible term that it is working.*
- *Let them know your expectation. Set expectation like what do you need to look out for when you write or speak etc.*
- *Need to give students 'wait' time.*
- *Do a demonstration.*
- *Show them what is a good journal, how you would write, define, and then explain, illustrate and give examples.*
- *Give a model answer, so students will know what is expected. When they see another person actually writing it down, so they will think they can do it too.*
- *Go down to the individual students, show them how their writing can be improved by showing them their classmate's work and ask them to make a comparison and then ask them how they think they can improve their own writing.*
- *Teach them how to present something orally, what you would say, how to say and why you say this, how to do a proper conclusion, and how to behave yourself in front of the audience.*

- *It takes a change in the students' mindset, encourage them, do more often, and practice more.*

Views from the students. Similarly, students were also asked what kind of strategies they had adopted to overcome their initial difficulties when working on the new tasks, and how teachers could have done to help them work on the new tasks more efficiently.

The followings were their comments:

- *Take reference… see how my friends write and then I'll improvise.*
- *Teacher gives an example.*
- *Teacher demonstrates.*
- *Teacher guides us and teaches us where to start.*
- *Easier to understand now rather than the first time we are doing it.*
- *Let us do more oral presentations, by encouraging us to speak more.*
- *During oral presentation, she will start a sentence for us to continue so that we know how to present…*
- *Work with partners and discuss.*
- *She gives hints.*
- *I ask my friends, then they show me some of the methods, and from there I try to do my own.*
- *Let them have group work.*
- *Play games that require writing.*
- *Teacher demonstrates more and don't be so strict lah!*
- *Encouragement.*
- *Give more examples, references.*
- *Give them tips.*
- *Give something like an example we can follow the example.*

Comments and suggestions for better integration of the new tasks
When students were asked to give comments and suggestions for better integration of the new tasks in mathematics classrooms, the usual respond received were that they did not really understand the questions. Even when some students did try to answer the questions, the answers were usually vague and unclear. Therefore in this section the report would only describe the results of the responses given by the teachers.

The teachers claimed that having a good, practical resource book for teacher's referencing was most important and essential for teachers to begin using the new tasks for the very first time. One teacher gave her idea of a resource book as: *"One which you can open up and say: ok today I want to use this; rather than starting from scratch in designing the task. Giving teachers ideas and then let them modify or make changes to the tasks to suit their individual's teaching."* The resource book should include samples of different kind of journal writing tasks and oral presentation tasks with samples of students' work. The samples of the students' work could be shown to students as 'model' work or use them for illustration or demonstration purposes. The resource book should also contain sample rubrics and how these sample rubrics can be used to assess students' work.

In addition, the teachers expressed that providing training and workshop for teachers who wished to integrate these new tasks in their classroom could also be helpful. In the training or workshop, the teachers would like to see what had been done or experienced by other teachers who had knowledge and know-how to use and integrate the new tasks appropriately and efficiently in their own classroom teaching. The teachers interviewed also suggested providing video clips showing some authentic classroom examples in training and workshop sessions. Other suggestions included having a session to help teacher-participants to create rubrics and the use of rubrics to do markings, and how to modify the new tasks to efficiently integrate them into the scheme of work.

Moreover, all the teachers had recommended that the integration of new tasks be implemented to, for example, all mathematics teachers in the school so that teachers as a department can come together to share experiences, discuss, collaborate, and to provide comment and feedback that would have an important effect on their daily teaching. Specifically, one teacher mentioned this: *"…if it's like a department thing, every teacher doing, we can actually collaborate as a department. So if it is just me alone, coming up with the task on my own and you don't know how the other classes doing, so it is quite difficult…But if it is a department doing it together, I think it will be easier."*

One other comment made by the teachers was that they felt that the assessment for mathematics should also change to accommodate the smooth integration of the new tasks, or else both teachers and students would not be able to see the needs to do such tasks. One teacher especially noted that the mathematics questions in the school examination and the national examination had always been the same, if teachers were encouraged to use the new tasks, they may question for example: *"Why do we have to do it when the national exam is not changing much?"* Consistently, another teacher also expressed that: *"The assessment work has to change"*, such that the assessment and the work that students do in the daily learning should be parallel. This teacher gave an example stating that although sometimes it could be very difficult to grade oral tasks but teacher's observations and statements written about students' competence in doing the oral tasks were important information and this information should be credited as part of the students' mathematics assessment.

Summary

The teachers and students interviews were conducted to gather direct and important information about the use of communication tasks. Three broad categories of questions were posed to the teachers and students in the experimental classes. The questions were designed to find out about (a) both teachers' and students' own experience and understanding about journal writing and oral presentation, (b) teachers' opinion and comments about their students' ability to perform journal writing and oral presentation tasks; similarly students were also asked to comment about their capability in doing these new tasks, and (c) teachers' comments and suggestions for effective and efficient implementation of journal writing and oral presentation tasks during daily teaching.

The results revealed that generally, both teachers and students were able to identify the differences between the conventional tasks and the communication tasks used in this study. The common understanding were that compared to the conventional tasks which were usually focused on getting the right answer or solution, communication tasks were non-routine; they required substantial explanation, description, reasoning, illustration, and systematic and coherence presentation. Both teachers and students also believed the importance of developing these communication skills in mathematics learning, and that the use of journal writing and oral presentation tasks were deemed feasible to be implemented and integrated in the daily lessons only if appropriate frequency was applied. Teachers also believed that the communication tasks would benefit students' learning because these new tasks provided an avenue for students to engage in in-depth understanding when learning mathematics.

The main difficulty that the teachers had was how to integrate the new tasks into their daily curriculum that was already pre-planned in the scheme of work. Teachers were guided, supervised and made to see that the new tasks were not supposed to be add-ons but treated as part of what students did in their normal daily lessons. Although teachers did not express a lot of difficulties in constructing the journal writing or oral presentation prompts, the new tasks were relatively new to them so they took a long time to think and reflect upon the prompts before implementing them into their daily teaching instructions. They also took a lot of time in reading students' journal writing. Thus, they conveyed that giving students quick and effective feedback on their writings were challenging for them. In addition, giving effective and constructive feedback to students also meant more instructional time in class needed. Thus the time factor was another challenge for teachers as well. Teachers had to struggle in completing the pre-planned and pre-existing scheme of work, at the same time trying to learn to implement and integrate the new tasks in their teaching.

The main difficulties that students faced when doing the new tasks were that they did not know how to start and what were expected of them. They were concerned that their command of the language could either hinder or advance their performance in doing communication tasks. However, they also expressed that with consistent practice and appropriate guidance and encouragement, they believed that they would do better.

Teachers believed that their students were generally capable of doing communication tasks but we should not expect them to be able to them well right from the start. Students must be given the chance to make mistakes, and talk and reflect about their work. Students' attitudes toward learning, not necessarily be related to the new tasks must also be improved in order for them to learn better.

Teachers gave their comments and suggestions toward a more effective and efficient way of implementing the new tasks in the daily mathematics teaching. Teachers expressed that first of all, in order to do this well, they should be supported, guided, and trained. Second, because these are new tasks, which meant new things for both teachers and students, more time should be given to them to explore and try out; do not expect both teachers and students to do them right and well. Third, since these new tasks are going to be seen as another kind of 'assessment', then how should we report the results, how much do we use these results to replace some of the current assessments, and how should we change our current assessments in order to accommodate the new assessment mode, are questions that need to be considered carefully.

9.5 Conclusions, Implications, and Recommendations

9.5.1 Summary and conclusions

This sub-study was part of a large research project on integrating new assessment strategies into daily mathematics classrooms teaching and learning. This report focused on the use communication tasks as one of the new assessment strategies in two Singapore secondary schools. There were four participating secondary one mathematics classes that received communication tasks for about 18 months and these students were tracked until they reached secondary two levels. There were also four corresponding comparison classes whose students and teachers did not receive any communication tasks and they went about in the mathematics lessons in the usual normal ways. More specifically, this study had employed the use of journal writing and oral presentation as the communication tasks in the participating experimental classes. Thus, this study was intended to investigate the effects of using journal writing and oral presentation in Singapore secondary school students' daily mathematics learning, and how these new strategies can be effectively and efficiently integrated into the daily mathematics classrooms teaching.

There were basically four main types of data collected in this study. The first were students' responses to the pre- and post-questionnaire surveys. Results found that generally as students were promoted from secondary one level to secondary two level, their general attitudes toward the subject mathematics and their own perceptions about their mathematics learning had become more negative. This phenomenon was observed in all the students, students in the experimental classes as well as in the comparison classes. The students in the experimental classes received communication tasks in their daily mathematics learning for about 18 months. 31 question items were constructed to find out about their general attitudes toward the use of these new tasks. The results found that generally all the students did not like to do these tasks very often or like to see more of these new tasks in their mathematics classrooms! However, they did belief that these new tasks could benefit in their learning of mathematics. Interviews with them further confirmed that they actually did not mind having to perform these new tasks in the mathematics classroom but they would like to be further guided and encouraged. This result was quite understandable because students were new to these tasks and 18 months of intervention was believed not long enough to make a positive or great impact on their attitudes toward either their general mathematics learning or toward their general perceptions of the new tasks. The study believed that the impact of using communication tasks on students' attitudes toward mathematics and mathematics learning should come gradually, and moreover attitudes were believed to be affected by many other factors for examples teacher's beliefs or views about them, teacher's teaching etc.

The second main data collected were students' school-based mathematics examination scores and they were M2004, F2004 and M2005. Results found that the exposure to communication tasks did not at least have any negative effect on students' mathematics scores on conventional tasks. There was no evidence to show that students' exposure to communication tasks had caused them to do badly in their conventional mathematics examinations. By M2005, the average number of interventions done by each experimental class was about 16. This meant had made an impact on students' results for M2005 as the results revealed that all the experimental classes compared to their corresponding comparison classes had significantly better mathematics scores for M2005 except for the Normal Academic class in the high-performing school where no significance was detected. Although no significance was detected, their scores were increased from F2004 to M2005. This particular class was observed to have the least seriousness in their doing of the intervention tasks throughout the intervention period. Teacher's interviews further confirmed that this class was particularly less interested in anything that they do, whether it was pertaining to the usual conventional task or the intervention tasks.

The third main data collected were students' achievement scores on the unconventional "journal writing task" tests. The main focus about these tests were to measure two aspects of students' learning namely (a) students' mathematical reasoning, and (b) students' coherence presentation skills. Results found that all the experimental classes did significantly better than their corresponding comparison classes in both the aforementioned aspects. Results seemed to indicate that students who were exposed to communication tasks improved in their mathematical reasoning and presentation skills.

The other main data collected were both teachers and students interviews. The interviews were intended to obtain direct information and confirmation from both teachers and students issues pertaining to the use of communication tasks in teaching and learning of mathematics. Results found that teachers and students were able to identify the differences between the communication tasks and the usual conventional work that students did in the daily lessons. Generally both teachers and students believed that doing communication tasks would help and benefit in students' mathematics learning. However, the command of the English language was also believed would either hinder or benefit students' abilities to perform in the communication tasks. While both teachers and students had expressed their willingness to use these new tasks in their teaching and learning, students had especially made known that the frequency of doing these new tasks should be gradual. Students also expressed that they wanted more guidance and encouragement from their teachers. On the other hand, teachers also needed time and guidance to be more familiar with the new tasks.

9.5.2 Implications and recommendations

First of all, teachers need to recognize that giving opportunity for students to be involved in active and meaningful written and verbal communication in mathematics classrooms is an essential process for their learning and knowledge acquisition. Teachers must believe that practicing good writing and oral presentation should be made as part of the mathematics learning activity. Although learning of mathematics has often been conceived as merely acquisition of computational skills and procedures, developing students' communicational skills is also believed to have an impact on students' reasoning and understanding which may ultimately affect their general achievement in mathematics. Thus, the emphasis of using communication tasks in mathematics classrooms for students' learning should be encouraged.

Similarly, students who probably have not been exposed to the emphasis of using communication tasks in their mathematics learning during their past schooling years need to understand why developing their communication skills are important in their mathematics learning now. In the same way, students should be informed clearly about the rationale and expectation of working on the new tasks in their mathematics classrooms. Teachers may need to encourage and help change students' mind-set and their attitudes toward the willingness to accept doing communication tasks during mathematics learning.

The researchers have the following recommendations for teachers: First, it has to be made known to students what teacher would like students to achieve at the end of the day through working on communication tasks, remembering that these tasks are generally new tasks for them and therefore teacher may need to begin by teaching some specific skills in order to help them work on the new tasks.

Second, teachers must not expect 'perfect' results from students right from the beginning. Allowing wait time for students to think about their work and chances for them to make mistakes are parts of students' learning processes. Third, teachers should start "small", that is to say, students may start with a short new task and work on it gradually once every two weeks, then subsequently increasing the frequency and the length of the task, bearing in mind the time constraint etc.

Last, teachers must give consistent and constructive feedback about students' work so that they can further work on their weaknesses and improve in their performance.

The results from the study suggested that teachers were generally competent in using communication tasks, and students were generally capable in doing these new tasks after a period of time. Integrating communication tasks into daily mathematics classrooms teaching and learning is seen feasible but it has to be done gradually and systematically. Teachers need to recognize that using communication tasks for the first time in their instructions can be quite challenging, thus the idea is to start small; do things a little at a time.

In this connection, the researchers have the following suggestions to recommend. When using journal writing, teachers may want to start students off with short writing tasks and subsequently increasing the number of prompts in the journal writing task asking students for more substantial explanation or reasoning, and journal writing task can be used once every two weeks.

As for oral presentation tasks, we recommend that pre-planned oral presentation task be done progressively throughout the school term, beginning with once or twice per school term, and impromptu oral task can be used consistently everyday asking students to summarize the day's lesson for example. Although teachers did expressed that it was time consuming for them to constantly give feedback on students' work on the new tasks, but when used appropriately, these new tasks can still be implemented in our daily classrooms.

The important questions to ask oneself are: (i) For what purposes do I use these tasks today? (ii) Am I able to use these tasks to elicit students' learning and understanding of mathematics that I want to find out? Given that when students are getting into a new stage of learning, they are usually more open and acceptable to new ways and styles, communication tasks should be implemented and integrated at the beginning of the secondary education. So that by the end of their four years of secondary schooling, students would have sufficient amount of exposure and opportunities to work on communication tasks.

Similarly, the results from the study also suggested that students were generally more competent in doing journal writing than oral presentation. Students' attitudes and feelings seem to suggest that writing is less 'threatening' compared to speaking in front of the whole class. Teachers however believed this phenomenon had to do with students' competence in using their verbal language and students' general self-confidence. The learning of mathematics has always been associated with learning of solving mathematics problems using mathematical formulae, strategies, logarithmic functions, etc as long as solutions are mathematically sound and correct. Very often there is no necessity for students to explain themselves why and how they solve mathematical problems using their own words and thoughts. Thus, students are not used to verbalize their mathematics understanding in words. Researchers suggest that before the formal implementation of communication tasks in mathematics classrooms, teachers should help students prepare themselves by doing the followings:

- Go through with students the points and steps to take note when writing and when doing a presentation.
- Show students samples of writing that demonstrate 'good', 'average' or 'bad' writings by pointing out exactly what are the criteria that makes a piece of writing 'good' and what are the necessary information needed for a piece of writing to meet expectation, etc.
- Give students chances to speak and do some hands-on with whole class evaluation, creating a non-threatening situation allowing friendly critiques and discussion.

Given the fact that communication tasks are new things to be done in mathematics classrooms teaching for instruction, the researchers recommend that careful planning and consideration be specified at the school level to ensure that smooth implementation can be carried out. School administrators and educators need to collaborate and build proper groundwork or foundation that will support the new tasks to be integrated into the mathematics curriculum. For example, from the interview data, teachers had indicated that providing training, preparation workshop or some form of professional development lecture that will aid and support them in the implementation of the new tasks in their classrooms is vital.

Some of the other consideration preceding the actual implementation should also be taken into account. For instance, all teachers from the mathematics department may need to spend time to learn more about what are communication tasks, have a general consensus about why the department beliefs that communication tasks would be used in mathematics teaching, how communication tasks could be used during classroom instruction, how communication tasks could be constructed such that these tasks are integrated in the mathematics syllabuses, and how information gotten from students' work on communication tasks be used. Thereafter, the department may need to devise a plan to evaluate the outcome of the implementation so as to record the results of all the happenings in the individual classroom. This is necessary because both teachers' and students' difficulties of using the new tasks may surface as a result from the evaluation and therefore developing appropriate plans for supporting teachers and students would have to be done.

In order to investigate more accurately whether communication tasks ultimately could change students' attitudes toward mathematics and mathematics learning, or have an impact on students' general achievement in mathematics, we need to study for a longer period of time. Longitudinal studies may be needed to find out more about the results of students' general attitudes and whether students' abilities to perform mathematics have improved.

Moreover the communication tasks that are described in this study are relatively new ideas and strategies for both teachers and students, and therefore effective and efficient use of these new tasks would take a long time through constant evaluation, fine-tuning and modification. This is going to be a long term process and for that reason, further exploration and trial period are necessary to accurately document the success or unproductive episodes of the use of communication tasks in the daily classroom teaching and learning.

Chapter 10 Results and Findings (VIII): Student Self-Assessment (Secondary)[16]

10.1 Introduction

Student self-assessment is one of the four types of new assessment strategies that the Mathematics Assessment Project (MAP) investigated. The assessment requires students to do self-evaluation and self-reflection on their learning. With information about students' learning collected through students' self-assessment, teachers can understand their students' learning better and hence improve their instructions in classrooms, while students can also improve their learning from their own self-assessment activities. In a large sense, self-assessment is an assessment mainly used for improving students' learning rather than for the measurement of students' performance.
This chapter reports the results and findings of the MAP sub-study focusing on student self-assessment at the secondary school level.

10.2 Research Question & Conceptual Framework

10.2.1 Research questions

Consistent with other sub-studies of the MAP project, the main research questions for the sub-study can be described as follows:

1. What are the influences of using self-assessment strategies on students' learning of Mathematics in their cognitive domain?

2. What are the influences of using self-assessment strategies on students' learning of Mathematics in their affective domain?

3. How can the self-assessment strategy be effectively integrated into the mathematics classrooms in Singapore?

10.2.2 Conceptual framework

One of the main reasons for using self-assessment is to let students be more responsible for their learning. By doing so, we hope that the students can be better motivated and be placed in their intrinsic flow of learning. It is a student-centred assessment.

The self-assessment activities in the study are further divided into three sub-types: self-evaluation, self-reflection, and self-construction. Self-evaluation asks students to evaluate their own learning particularly at the end of a period of learning (e.g., a chapter or a mathematics topics), self-reflection asks students to reflect on their own learning process participating their problem-solving process, and self-construction asks students to pose or construct mathematics problems based on what their have learnt over a period of time or topics.

In addition, in the process of conducting the self assessment a "Ten scenarios & prompts" was designed and used to engage students during their mathematics lessons. The flowchart below (Figure 10.1) shows the components of the self-assessment used in this sub-study.

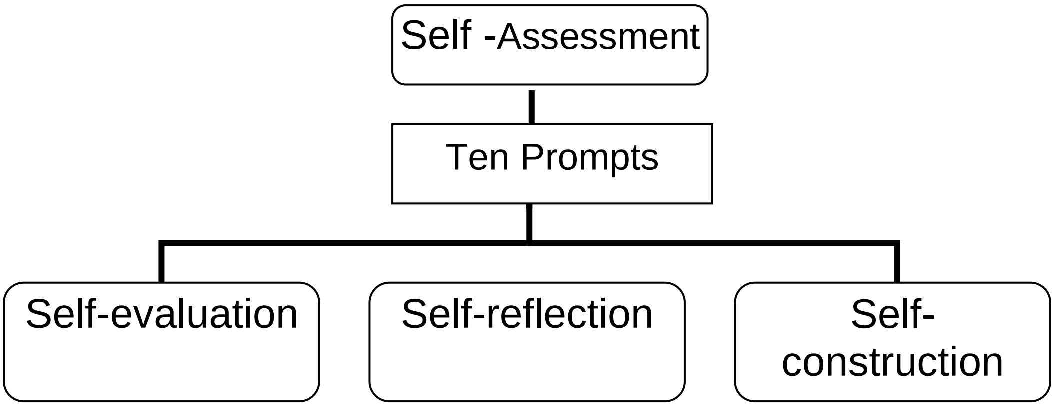


Figure 10.1. Components of the self-assessment in the sub-study

In short, students engaged in self-assessment activities are asked to evaluate their own learning through questionnaires, surveys, teachers' in-class prompts and questions, and the teachers react and make instructional decisions according to the information that they obtained from students' self-assessment. Through self-assessment, we hope that students will be able to perform a related task in a different environment and teachers can understand their teaching and students' learning in a more effective way.

10.3 Methods

10.3.1 Participants

This sub-study was carried out in two secondary schools. One of the schools is a high-performing school and the other is a non-high-performing school. The main criterion of ranking the schools as high-performing and non-high-performing was based on the year 1999-2002 GCE "O" Level Examination results released by the Ministry of Education (MOE). The high-performing school, School A, was randomly selected from the 50 best performing secondary schools in term of the average students' achievements, and the other non-high-performing school, School B, was randomly selected from the reaming schools.

In each school, four classes chosen for the study were recommended by the school which acted independently without interference from the researchers. Altogether eight classes were studied in two schools.

Furthermore, two of the classes in each participating school were selected as the experimental classes and the other two were selected as the comparison classes. Of the two experimental classes, one was from the Express Stream and the other from the Normal Academic Stream. Equivalent classes from each stream were chosen for comparison purpose. During the 18 months of study, due to a variety of practical reasons there were some changes in both the student and teacher participants. Detailed changes are illustrated below.

Changes in students. The main change of students took place during the students' promotion of their learning from Secondary One to Secondary Two.

Table 10.1 summarizes the changes in participating students over the period of the invention. For convenience, the following notations for the experimental and comparison classes are used in the table.

1EH(A)	Sec 1 Experimental class of Express stream in School A	1EH(B)	Sec 1 Experimental class of Express stream in School B
1CH(A)	Sec 1 Comparison class of Express stream in School A	1CH(B)	Sec 1 Comparison class of Express stream in School B
1EL(A)	Sec 1 Experimental class of Normal Academic stream in School A	1EL(B)	Sec 2 Experimental class of Normal Academic stream in School B
1CL(A)	Sec 1 Comparison class of Normal Academic stream in School A	1CL(B)	Sec 2 Comparison class of Normal Academic stream in School B
2EH(A)	Sec 2 Experimental class of Express stream in School A	2EH(B)	Sec 2 Experimental class of Express stream in School B
2CH(A)	Sec 2 Comparison class of Express stream in School A	2CH(B)	Sec 2 Comparison class of Express stream in School B
2EL(A)	Sec 2 Experimental class of Normal Academic stream in School A	2EL(B)	Sec 2 Experimental class of Normal Academic stream in School B
2CL(A)	Sec 2 Comparison class of Normal Academic stream in School A	2CL(B)	Sec 2Comparison class of Normal Academic stream in School B

Table 10.1. Changes in student participants

Classes	School A		School B	
	Experimental	Comparison	Experimental	Comparison
Express 2004	1EH(A) 40 students	1CH(A) 39 students	1EH(B) 41 students	1CH(B) 41 students
Express 2005	2EH(A) 39 students from 1EH(A)	2CH(A) 39 students, all from 1CH(A)	2EH(B) 18 from 1EH(B) 12 from 1CH(B) 3 from 1E3 2 from 1E4 2 from 1CH(B) Total 37 students	2CH(B) 22 from 1CH(B) 12 from 1EH(B) 2 from 2E3 1 from 1E4 3 transferred in Total 40 students
Normal Academic 2004	1EL(A) 41 students	1CL(A) 40 students	1EL(B) 39 students	1CL(B) 38 students
Normal Academic 2005	2EL(A) 41 from 1EL(A) 4 from 2T1 Total 45 students	2CL(A) 40 from 1CL(A) 3 from 2T1 Total 43 students	2EL(B) 19 from 1CL(B) 11 from 1EL(B) 3 from 1A3 3 from 1E3 3 from 1E4 4 from 1CH(B) Total 44 students	2CL(B) 18 from 1EL(B) 19 from 1A3 Total 37 students

In School A, one student in the experimental Express class left the school at the end of 2004 and four students from the Normal Technical stream joined the experimental Normal Academic class in 2005. The students were the same throughout the intervention period for the comparison class in the Express stream. Three students from the Normal Technical stream joined the comparison class in the Normal Academic stream. Although there were changes of students in 2005, the students profile for the experimental classes and the comparison classes were considered similar in 2005 with realistic and minimal changes in Secondary School A.

In School B, the changes had affected all the experimental and comparison classes. In the Express stream, only 18 out of 38 in the 2EH(B) were from the previous experimental class 1EH(B). Only 12 out of 40 in the 2CH(B) were from 1EH(B). Similarly, from the table that indicates the movement of students from 2004 to 2005, 1EL(B) pupils (2004 experimental class) have been distributed into 3 classes. 2EL(B)(by Teacher M) has the 11 of them, 2A2 has 10 of them and 2CL(B)(by Teacher N) has 18 of them. The rest are in 2A2. At first glance, 2CL(B) which has the highest number of students from 1EL(B) (2004 experimental class) was thought to be best for the experimental class in 2005. However, it was found later that 2CL(B) could not be the experimental class in 2005 due to two main reasons.

Firstly, the mathematics teacher of 2CL(B) (Teacher N) was an untrained graduate teacher waiting to join National Institute of Education in May 05. He had injured his leg and was on medical leave for a period of time which was unknown to the researchers during that decision making period.

Secondly, while Teacher N was on medical leave, the class was taught by a relief teacher, Teacher R, who was a Junior college student who had just completed her ‘A’ level examination. As a relief teacher, Teacher R could leave the service at anytime. This had created another uncertainty for the researchers. The two uncertainties, one being the duration of the medical leave that Teacher N will take and two being how long Teacher R will stay to teach 2CL(B), had caused much worry among the researchers.

These two uncertainties would be too stressful as well as too challenging, if not impossible, for the relief teacher or the contract teacher to carry out the intervention in 2CL(B). Due to these uncertainties, 2EL(B) was suggested to be the experimental class for the Normal Academic stream in 2005. Teacher M, the mathematics teacher of 2EL(B) was willing to participate in the project and agreed to carry out the interventions for his class.

As the involvement for the project had been the students and teachers of the 2EL(B) and 2CL(B) for several weeks, in order not to disturb the school teachers and students too drastically, a swap was made. 2CL(B) hence became the comparison class for the Normal Academic stream in School B in 2005. Under these circumstances, 2A2 had neither been considered as the experimental class nor the comparison class.

Changes in teachers. Besides the changes in students, changes in teachers also added the challenges to the data collection.

Table 10.2 below shows the participating teachers' profiles and the changes of teachers for the participating schools for both the participating classes.

School A had minimal changes in the teaching staff and all mathematics teachers hold mathematics degrees. School B has more changes in the teaching staff and mathematics teachers hold various academic degrees. Among all the mathematics teachers, only one of the subject teachers had a mathematics degree, others hold degrees in Chemistry, Engineering, Accountancy or Business Administration.

Table 10.2. Changes in teacher participants

	School A		School B	
	Experimental	Comparison	Experimental	Comparison
Express 2004	1EH(A)	1CH(A)	1EH(B)	1CH(B)
Teacher (Duration; length of teaching experience)	Teacher Y (Jan 04-Dec 04; 6 yrs)	Teacher Z (Jan 04-Dec 04; 29 yrs)	Teacher J (Jan 04-Jun04; 3yrs) Teacher G (July04-Dec 04; 10 yrs)	Teacher W (Jan 04-Jun04; 3 yrs) Teacher M (July 04-Sep 04; 0.5 yrs) Teacher D (Oct 04-Dec 04; 2 yrs)
Express 2005	2EH(A)	2CH(A)	2EH(B)	2CH(B)
Teacher (Duration; length of teaching experience)	Teacher Y (Jan 05-Dec 05; 7 yrs) (with 2 months taken by a trainee teacher)	Teacher Z (Jan 05-Dec 05; 30 yrs)	Teacher C (Jan 05-April05; 1.5 yrs) Teacher T (UGT) (May 05-Jun05) Teacher C (July 05 –Dec 05)	Teacher D (Jan 05-Dec 05; 2 yrs)
NA 2004	1EL(A)	1CL(A)	1EL(B)	1CL(B)
Teacher (Duration; length of teaching experience	Teacher K (Jan 04-Dec 04; 25 yrs)	Teacher K (Jan 04-Dec 04; 25 yrs)	Teacher A (Jan 04-Jun04; 0.5 yrs) Teacher B (July04-Dec 04; 0.5 yrs)	Teacher A (Jan 04-Jun04; 0.5 yrs) Teacher C (July04-Dec 04; 0.5 yrs)
NA 2005	2EL(A)	2CL(A)	2EL(B)	2CL(B)

Teacher (Duration; length of teaching experience	Teacher K (Jan 05-Dec 05; 26 yrs)	Teacher X (Jan 05-Dec 05; 14 yrs)	Teacher M (throughout; 1.5 yrs)	Teacher R (Jan05-Fed 05; relief teacher) Teacher N (Mar05-April05; contract teacher) Teacher O (May 05 –Dec 05; 5 yrs)

In School A, the teachers of the experimental classes and the comparison classes are not very comparable in terms of years of experience. However, if the experimental class teacher is a Head of Department, we consider her to be strong in her ability that may be comparable to the teacher teaching the comparison class. In the Normal Academic stream, we had the same mathematics teacher teaching both the experimental and comparison classes in 2004. Due to this undesirable condition, in 2005, the comparison class in the Normal Academic stream was taught by another teacher who has 14 years of teaching experience. The information we gathered so far showed that teacher X who taught the comparison class in 2005 is an experienced and strong teacher. As a mathematics teacher, teacher X is not weaker than teacher K who taught the experimental class. Even though there is a gap in the number of years of teaching experience, according to the evidence that the researchers gathered, teacher X is by no means a weaker teacher if not stronger than teacher K.

In School B, there are many changes in teachers almost every semester. There are even more changes in 2CL(B) in the first semester. However, when we studied in detail, we found that from Jan 2004 to May 2004, the school has comparable teachers teaching both the experimental class and the comparison class for the Express stream (both classes were taught by teachers with 3 years teaching experience) and the same teacher teaching both the experimental class and the comparison class for the Normal Academic stream. However, from Jun 2004 to Nov 2004, Express stream has a very experienced teacher (10 years) teaching the experimental class and a less experienced teacher (2 years) teaching the comparison class. The school had arranged two comparable teachers to teach the experimental class and the comparison class for the Normal Academic stream (both classes were taught by teachers with 0.5 year's experience).

10.3.2 Instruments

Data in this study were collected through a variety of ways. This section describes the instruments we designed and used for data collection.

Questionnaire surveys (Pre- & Post-)
A survey questionnaire with part (a) on students learning attitude was conducted before the intervention and after the intervention, as mentioned in Chapter 2 (also see below). Part (b) of the survey was to find out the ways mathematics lessons were conducted and how often the students had been exposed to the self-assessment tasks. Both part (a) and part (b) of the survey was conducted before the interventions to all experimental classes and comparison classes of both schools. After the intervention period, only part (a) of the survey was given to all the experimental classes and the comparison classes of the schools. Part (b) of the survey however was given to the comparison classes only as it served no purpose to the experimental classes after intervention. The objective of the pre-intervention survey was to provide baseline information on both attitudes and on students' ability in Mathematics. The post-intervention survey was conducted 18 months later at the end of the experimental period. Post-intervention surveys were designed to be parallel to the pre-tests. The questionnaire items on self-assessment can be found in Section A of Appendix 10.1.

Part (a): Questionnaires on attitude. Part (a) has 22 questions and part (b) has 9 questions. Part (a) questions are to find out the participants' attitude in terms of their general attitude towards mathematics (G), their beliefs in learning mathematics (B), their feelings towards their achievement in mathematics (A), and their attitude towards their performance in the learning of mathematics (P). 12 of the questions are asked in the positive way while the other 10 are asked in the negative way to ensure each participant has a consistent response to the question asked.

The questions about students' general attitude (G) in the learning of mathematics are:
Q1: I enjoy doing mathematics,
Q5: (reverse) Mathematics is hard for me,
Q9: Mathematics is interesting to me,
Q13: (reverse) I don't have good feelings about mathematics,
Q16: I like spending time on studying mathematics and
Q19: (reverse) I don't like to attend math lessons.

The questions about students' beliefs (B) in the learning of mathematics are:
Q4: I believe mathematics is useful,
Q8: It is important to know mathematics nowadays,
Q12: (reverse) Studying mathematics is a waste of time, and
Q18: I will use mathematics a lot as an adult.

The questions about students' feeling towards their achievement (A) in the learning of mathematics are:
Q2: I am never under a terrible strain in a math class,
Q6: I am not afraid of doing mathematics,
Q10: (reverse) I am unable to think clearly when doing mathematics,
Q14: (reverse) I feel lost when trying to solve math problems,
Q17: (reverse) It makes me nervous to even think about having to do a math problem.
Q20: I have a lot of confidence when it comes to mathematics.

The questions about students' attitude towards their performance (P) in the learning of mathematics are:
Q3: I am sure I can learn mathematics well,
Q7: I can get good grades in mathematics,
Q11: (reverse) I am not good at mathematics,
Q15: (reverse) I don't think I can do well in mathematics,
Q21: I like solving challenging math problems, and
Q22: (reverse) I would rather have someone give me the solution to a difficult math problem than to have to work it out for myself.

In short, Part (a) of the survey is mainly classified into four big areas with the questions distributed as shown below.

General attitude (G)	Belief (usefulness) (B)	Achievement/Anxiety (A)	Performance (P)
Q1,5,9,13,16,19 (6 Qns)	Q4,8,12,18 (4 Qns)	Q2,6,10,14,17,20 (6 Qns)	Q3,7,11,15,21,22 (6 Qns)

Part (b): Questionnaires on ways mathematics lessons were conducted. Only the comparison classes did the same part (b) as the pre-survey. Part (b) questions are to find out how the mathematics lessons were conducted. Part (b) mainly asked about the frequency students reflected on their learning of mathematics, how often the teacher encouraged alternative solutions, how often they constructed their own questions and how often they evaluated their own learning.

Though the classification of the survey questions can be subjective, it was carried out on the same group of students before and after the intervention period with both the experimental and the comparison classes. It has been classified evenly for all components of this Mathematics Assessment Project (MAP). This instrument aims to find out how often the students were using the alternative assessment and how much the students were aware of the self-assessment. This instrument would help the researchers to find out how comparable the comparison group is as compared to the experimental group in terms of the new strategy.

School-based examinations
Each student's results for all the 8 classes were tracked. Their school-examination performances were recorded. Since students took the same paper in the same PSLE examination, PSLE was used to compare across classes and schools. However, the school-based examination is common only for the same stream in the same school. It can therefore be used only to compare between classes of the same stream in the same school. The school-based results were therefore used to compare between the experimental class and the comparison class of each stream of each school.

Altogether, the researchers had collected three examination scores, namely Mid-year examination 2004, End-of-year examination 2004 and Mid-year examination 2005 for each student. The results were considered as an instrument for the researchers to find out whether the performance of the experimental class and the comparison class had any difference. Though students' examination performances did not vary exclusively based on the interventions, the results can be used as reference. Researchers are fully aware that many factors like teachers' experience, teachers' relationship with students, students' motivation, students' learning support and others might influence the outcome of students' examination results. Therefore no conclusive result was drawn based on the school results alone.

"Self-assessment" tests (Pre- & Post-)
A test on self assessment was carried out before and after the intervention period. This self-assessment test aims to find out whether students are familiar with the self-assessment task and how well they have cultivated the habits of self-assessment.

The test consists of two parts. Part A of the test has 20 questions on a five point scale questionnaire. Part B of the test has three problems. Part B for the Pre-test and Part B for the Post-test are two parallel tests. The test papers were constructed with consultation of the participating teachers. The first problem is to test the students' thinking and their decision making. This problem helps to differentiate the mathematical ability of students in problem solving. Problem two is divided into part (a) and (b). Both the parts require pupils to apply a similar mathematical skill – using combination skill for two similar situations in part (a) and part (b). Problem two will show how well students can self-evaluate and apply the same mathematical skill in different situations. It is also to find out whether students in experimental group will perform differently in pre and post intervention. Since they have been asked to do self-reflection and self-evaluate in the interventions, the researchers hope to find out whether the intervention has helped them develop the evaluation skills or transferring their mathematical skill to different situations? Question three, pupils are asked whether they can see the similarity in both part (a) and (b) with justification. This question hopes to find out whether students are better in reflecting on their learning after intervention and therefore, there is hope to see whether the interventions had changed the students' habits in learning mathematics. Copies of the tests can be also found in Appendix 10.1 (Section B1 for pre-test and Section B2 for post-test).

"Self-assessment" intervention tasks

Initially the researchers designed the templates of self-evaluation and self-reflection of two topics and the schools' teachers modified them to suit their students and their lessons. After a few rounds of interventions, the teachers started to design their own surveys and worksheets. The surveys were mainly to find out whether students have difficulties in understanding their lessons and whether they have anything to clarify with their teachers. At the same time students consolidated their lessons and did their reflection about their learning. The intervention worksheets were designed to find out the misconceptions and learning difficulties that students would like to further enhance or clarify. Samples of the self-evaluation and self-reflection worksheets are shown in Appendices 10.2 and 10.3 respectively.

Teachers gave the evaluation survey after each topic or chapter and reflection worksheet once a month. Some classes had tried the self construction test items and some had not.

These intervention instruments were used to build up students' self-reflection and self-evaluation habits. During the intervention period, the researchers helped the teachers summarize the evaluations and reflections. The summary though giving the researchers much insight of the intervention, usually did not give timely feedback to the students or the teachers. In fact, teachers made sure that they read through the students reflections before they went for their next lesson. According to the teachers, this was to ensure that they improved on their teaching instructions and fixed the problems occurred during the previous lesson.

For School A, 9 interventions were done in the Express stream for the 18 months (3 from Jan 04 to Jun 04, 4 from Jul 04 to Nov 04, and 5 from Jan 05 to May 05), and 6 interventions were carried out in the Normal Academic stream (1 from Jan 04 to Jun 04, 2 from Jul 04 to Nov 04, and 3 from Jan 05 to May 05). For the School B, 20 interventions were carried out in the Express stream (3 from Jan 04 to Jun 04, 7 from Jul 04 to Nov 04, and 10 from Jan 05 to May 05), and 22 interventions were carried out in the Normal Academic stream (3 from Jan04 to Jun 04, 8 from Jul 04 to Nov 04, and 11 from Jan 05 to May 05).

Classroom observation (including video recording)

Altogether, 66 classroom observations pertaining to self assessment were conducted. Out of the 66, 36 were done in School A and 30 in School B. Throughout the intervention period, in School A, 18 observations were done with the experimental class of the Express stream and 18 observations with the experimental Normal Academic stream. For School B, both the experimental Express stream and the experimental Normal Academic stream were observed 15 times. 5 video recordings were done in School A, 3 with the experimental Express stream and 2 with the experimental Normal Academic stream.

The classroom observation was employed as an instrument in this project for the researchers to find out how teachers integrate the self-assessment into their mathematics lesson. This instrument would give the researchers much insight and cue for the research questions.

Video Recording was another instrument that was used to remind researchers of the classroom discourse and as evidences for the research. The video recording was done after several classroom observations. Teachers and students were comfortable with it.

Interviews

Several interview sessions with the students and teachers of the experimental classes were conducted. As discussed in Chapter 2, interview questions can be divided into two main categories (See Appendices 2.4 and 2.5). Category one is about students' or teachers' own experience and understanding about self assessment. Category two is about students' or teachers' opinions or suggestions on the use of self-assessment.

Since the self-construction test was carried out only in School A, the students and teachers of the experimental classes were also interviewed about their experiences and views on the self-construction test**.**

Interview with teachers. For School A, the teachers in the experimental classes had taken part in the project for 18 months. They were interviewed individually by two researchers on an afternoon. Their views and experiences were recorded and were transcribed. For School B, only the mathematics teachers of the Sec 2 experimental classes were interviewed. Both the teachers had taken part in the project for the final 6 months.

Due to time constraint and the busy schedule in the school, the other participating school teachers were not interviewed. As a matter of fact, through informal conversation with the Heads of Mathematics Department and leaders of the schools, feedback regarding self-assessment was collected informally. For school A, they implemented the self-assessment to their Sec 1 cohort in 2005. This indication is a strong indication that teachers in the school welcome the self-assessment strategy and have a strong belief that the self-assessment will help their students in their learning of mathematics.

Interview questions for students. A number of students who participated in the project for 18 months were interviewed focusing on their experiences and feelings about doing self-assessment worksheets and surveys. They were asked to share their self-assessment experience and to suggest some constructive ways to improve the integration of the new assessment into classroom teaching.

The students interviewed were selected randomly by the school. Researchers set the condition that they would like to interview a sample of students who had participated in the project for 18 months. They were interviewed in threes. It was felt that this would be a more supportive environment for the students. It was believed that students who were in the project for 18 months would be able to give a more detailed and clearer description of their experience and feelings.

In addition, during class observations, field-notes were taken and later summary reports were written to update the other team members. Weekly updated information was circulated among the members and monthly meetings were conducted. All these are to keep team members informed of the progress of the research outcomes.

10.3.3 Procedures and data collection

Student self-assessment was demonstrated mainly through the following four tasks or activities, as mentioned earlier.

1. Student self-evaluation. As one of the participating teachers put it “Assessment and evaluation were often inter-related. Assessment refers to the process of gathering information about students’ abilities and using such information to decide on the future instruction. Evaluation, on the other hand, is the process of assigning value to students’ work.” This research focused more on summative evaluation where information after instruction was collected to summarize how students had performed and to determine grades at times.

To integrate student self-assessment into the teaching and learning process of the experimental classes, the students did self-assessment sheet, once a week or once a topic. A generic form was provided by the researchers and the teachers filled in the topics they are teaching and carried out the survey with their students. Once a topic or a chapter was completed, teachers would then ask students to do self-evaluation which focused on students’ cognitive domain and metacognition (monitoring their own thinking).

A sample of self-evaluation with students' responses is attached for reference in Appendix 10.2.

2. Student self-reflection. Student self-reflection is to be conducted as and when it is needed. It is a less visible and measurable ingredient of self-assessment. Teachers integrated the idea, by asking students to do more self-reflection while students frequently monitor their own thinking process, and be engaged in the stage of "looking back". Teachers took necessary instructional action based on the information they collected from students' self-reflection activities. A sample of self-reflection sheet is attached in Appendix 10.3

3. Student self-constructed test. Students were asked to construct questions after their revision. Questions constructed were amended by their teacher to form a test paper for the whole class. The self-constructed test was successfully carried out in School A on Chapter 9: Perimeter and Area of Simple Geometrical Figures and Chapter 10: Volume and Surface Area, while School B managed to get students to set questions but did not use the students' questions for their class test.

A survey and an interview were carried out to find out how students felt after the activities. Both the survey items and interview questions are attached in Appendices 10.4 and 10.5 respectively.

4. Ten prompts. The ten prompts guided teachers and students to have a deeper thinking and reflection in their teaching and learning. They are expected to be used during teachers' daily classroom discourse with students whenever appropriate and helpful. The prompts can be found in Appendix 4.1.

10.3.4 Limitations of the study

Like other classroom-based educational studies, there were many challenges and unexpected difficulties in this sub-study carried out in the two schools, which posed the limitations of the study.

Limitation of study in School A. Though the change of student participants in School A is small, we could not ignore the possible effect that may impinge on the significance of the study. The validity and reliability of the interventions could have been affected somehow. The changes in the teacher participants and the differences in the teachers' abilities in terms of teaching experience had caused much worry that the comparison class may not be really comparable. Due to many other constraints that School A had to face and unavoidable differences among the abilities of teachers and their teaching experience, the researchers have never ignored the teacher factors that will influence the students' performance.

Limitation of study in School B. There were many changes in School B. Both teacher and student participants were changed almost every six months. Due to the fact that the main changes happened when the student participants were re-distributed into several classes, the researchers hence would like to treat the interventions in School B as two different studies. The first intervention was from Jan 2004 to Dec 2004 where there were 11 interventions and the second was from Jan 2005 to May 2005 where there were another 11 interventions.

Because students in the experimental class and the non-experimental class in School B were mixed in Secondary two, no matter how the data were treated, there was still a possibility that the data were corrupted. Therefore, there was a stage that researchers wanted to discard the data totally, however, at a second look, we feel that the data still can be meaningful when we analyse it with all the constraints taken into consideration.

10.4 Research Findings

Below we shall report the results and finding of the sub-study. We start with the results from questionnaire surveys.

10.4.1 Questionnaires on students' attitudes

The pre-survey result of the questionnaire on attitudes in School A shows that the experimental class had significantly better attitude as compared with the comparison class in Express streams, but no significant difference was found in most of their attitude between the experimental class and the comparison class for the Normal Academic stream. The only attitude that shows significance difference between the experimental and comparison class in the Normal Academic stream is that more students enjoy doing mathematics in the experimental class than that of the comparison class.

For the Express stream, only two items do not show significant difference. They are the item that asks whether "they are able to think clearly when doing mathematics" and the item that says "I am sure I can learn mathematics well".

The pre-survey result of the questionnaire on attitude of School B shows that the experimental class is better than the comparison class for the Express stream and is significantly better in three items that say "I am never under a terrible stain in a math class", "I am sure I can learn math well" and "I am unable to think clearly when doing math". Though the pre-survey result shows the comparison class of the Normal Academic has a better attitude than the experimental class, there is no significant difference in the findings.

All the classes of both schools showed positive attitude towards Mathematics and learning of Mathematics with their averages more than 5 in the 9-point scale in the pre-survey. Post-survey has exactly the same questions as the pre-survey. Post survey result shows that when we compare within classes for School A, the following were observed:

- Students in experimental class for the Express stream provided more negative responses in the post- than pre-survey on all the items and on 9 items (Item 4, Item 7, Item 8, Item 9, Item 12, Item 18, Item 19, Item 20, Item 21) they provided significantly more negative responses in the post-survey; Similarly, students in comparison class for the Express stream provided more negative responses in the post- than pre-survey on all but 3 items (Item 2, Item 6, Item 19) and on 3 items (Item 4, Item 8, Item 12) they provided significantly more negative responses in the post-survey;

- Students in experimental class for the Normal Academic stream provided more negative responses in the post- than pre-survey on all but 3 items (Item 16, Item 18, Item 21) and on 1 item (Item 19) they provided significantly more negative responses in the post-survey; Similarly, students in comparison class for Normal Academic stream provided more negative responses in the post- than pre-survey on all but 6 items (Item 1, Item 4, Item 9, Item 17, Item 18, Item 19) but on 1 item (Item 1) they provided significantly positive responses in the post-survey;

Comparing within classes in School B, the following were observed:

- Students in experimental class for Express stream provided more negative responses in the post- than pre-survey on all but 9 items (Item 8, Item 13, Item 14, Item 15, Item 17, Item 18, Item 19, Item 21, Item 22) but no significant difference was detected; Similarly, students in comparison class for Express stream provided more negative responses in the post- than pre-survey on all but 1 item (Item 10) and on 11 items

(Item 1, Item 3, Item 4, Item 7, Item 8, Item 11, Item 12, Item 13, Item 15, Item 18, Item 19) they provided significantly more negative responses in the post-survey;

- Students in experimental class for Normal Academic stream provided more negative responses in the post- than pre-survey on all but 7 items (Item 5, Item 10, Item 11, Item 14, Item 16, Item 21, Item 22) and on 2 items (Item 1, Item 4) they provided significantly more negative responses and on 1 item (Item 21) they provided significantly more positive responses in the post-survey; Similarly, students in comparison class for Normal Academic stream provided more negative responses in the post- than pre-survey on all but 3 items (Item 16, Item 18, Item 21) and on 4 items (Item 10, Item 11, Item 12, Item 19) they provided significantly more negative responses in the post-survey;

- To express in mathematics notation we have the following EH(A)(Post with 6^+) < EH(A)(Pre with 13^+, 5 overlap) ($^+$significantly different) ;
  CH(A)(Post) < CH(A)(Pre)
  EL(A)(Post) < EL(A) (Pre) ;
  CL (A)(Post)< CL(A)(Pre, ex Q1)
  EH(B)(Post) < EH(B) (Pre) ;
  CH(B)(Post) < CH(B)(Pre)
  EL(B)(Post) < EL(B)(Pre) ;
  CL(B)(Post) < CL(B)(Pre)

The above show that the post-survey results in all classes do not show better attitude toward mathematics and learning of mathematics as compared to that of the pre-survey results. One possible reason could be due to some Recency Effect. The pre- survey was taken when school term just started and students were happy that they had completed their primary school education and were starting their secondary education which is new and full of hope to most of them. However, the post survey was taken when they had sat for their Sec 2 mid-year examination. To most of the students in secondary school, they faced more challenging mathematics problems and the syllabus is generally tougher, their examination results do not give them much happiness or joy to celebrate. There could be some apprehension and anxiety towards learning when they were doing the post survey which therefore resulted in a Recency Effect that made the post test result worse off. Further study was conducted to find out the changes in attitude between experimental classes and comparison classes.

Comparing between classes for School A

- In both the pre- and post-survey, students in experimental class for Express stream provided more positive responses than those in comparison class for Express stream; in particular, it was the case for all but 2 items (Item 5, Item 10) in the pre-survey and all but 1 item (Item 7) in the post-survey; moreover, students in experimental class for Express stream provided significantly more positive responses on 13 items (Item 1, Item 2, Item 4, Item 6, Item 9, Item 12, Item 13, Item 16, Item 17, Item 18, Item 19, Item 21, Item 22) in the pre-survey and on 5 items (Item 1, Item 8, Item 12, Item 13, Item 16) in the post-survey;

- In both the pre- and post-survey, students in experimental class for Normal Academic stream provided more negative responses than those in comparison class for normal academic; in particular, it was the case for 11 items in the pre-survey and 14 items in the post-survey; moreover, students in experimental class for Normal Academic provided significantly more positive responses on 1 items (Item 1) in the pre-survey and on 2 items (Item 12, Item 19) in the post-survey;

- In both the pre- and post-survey, students in experimental class for Express stream provided more positive responses than those in experimental class for Normal Academic stream on all the items; moreover, students in experimental class for Express stream provided significantly more positive responses on all but 2 items (Item3, Item10) and on all but 3 items (Item 7, Item 18, Item 20) in the post-survey.

Comparing between classes for School B

- In both the pre- and post-survey, students in experimental class for Express stream provided more positive responses than those in comparison class for Express stream; in particular, it was the case for all but 5 items (Item 12, Item 13, Item 14, Item 17, Item 21) in the pre-survey and all the items in the post-survey; moreover, while students in experimental class for Express stream provided significantly more positive responses on 3 items (Item 2, Item 3, Item 10) in the pre-survey and all but 5 items (Item 2, Item 6, Item 10, Item 12, Item 17) in the post-survey;

- In both the pre- and post-survey, students in experimental class for Normal Academic stream provided more negative responses than those in comparison class for Normal Academic stream ; in particular, it was the case for 13 items in the pre-survey and 12 items in the post-survey; while no significant difference between the two classes was detected in the pre-survey, students in experimental class for Normal Academic stream provided significantly more positive responses on 3 items (Item 10, Item 11, Item 14) and significantly more negative responses on 1 item (Item 1) in the post-survey;

- In both the pre- and post-survey, students in experimental class for Express stream provided more positive responses than those in experimental class for Normal Academic stream on all the items; moreover, students in experimental class for Express stream provided significantly more positive responses on 7 items (Item 2, Item 3, Item 5, Item 6, Item 7, Item 11, Item 12) in the pre-survey and all but 10 items (Item 1, Item 3, Item 4, Item 7, Item 8, Item 9, Item 13, Item 17, Item 19, Item 22) in the post-survey.

- In mathematics notation we have the following:
  $EH(A) > CH(A)$
  $EL(A) < CL(A)$
  $EH(A) > EL(A)$
  $EH(B) > CH(B)$
  $EL(B) < CL(B)$
  $EH(B) > EL(B)$

Generally, all the experimental classes do not change their relative position as compared to their comparison classes in both pre- and post- surveys. All experimental classes except the Normal Academic class in School A show a better attitude in both pre- and post- survey. The interventions do not change their relative positions, neither did the Normal Academic class in School A show any change in comparison with its comparison class. Though there are items like Q1: I enjoy doing mathematics, the experimental class is significantly more positive than comparison class in the Pre-survey. In the Post-survey, the experimental class is significantly most positive in items Q12: (reverse) Studying mathematics is a waste of time, and Q19: (reverse) I don't like to attend math lessons. Whether the reverse in attitude shed some light on the possible positive effect of intervention is still inconclusive. Nevertheless, it seems safe to say that the intervention had neutral or positive influence on the students' attitudes in mathematics learning, which is further supported in the data collected from interviews (see below)

10.4.2 Questionnaires on ways of teaching mathematics

Using Mode as the central tendency the survey results on the ways mathematics lesson were reported and analyzed. The results have shed some light on how the experimental classes and the comparison classes are exposed to the new strategies.

Pre-survey results for School A. For the Experimental Express class in School A-EH(A): the mathematics lesson were conducted without much use of the new strategies except for the items “teacher asks students to have more than one correct answers” and “teacher encourages students to solve mathematics questions in different ways”.(2 to 3 times a week).

For the Comparison Express class in School A-CH(A): the mathematics lesson were conducted without much use of the new strategies except for the items “teacher asks students to have more than one correct answers” and “teacher encourages students to solve mathematics questions in different ways”.(2 to 3 times a week).

For the Experimental Normal Academic class in School A-EL(A): the mathematics lessons were conducted without the use of the new strategies except for the items “teacher asks students to write down the reasons for their mathematics answers” and “teacher asks students to think about the reason for solving their mathematics problems”. Occasionally “teacher asks students to have more than one correct answer” and “teacher encourages students to solve mathematics questions in different ways” (2 to 3 times a week).

For the Comparison Normal Academic class in School A-CL(A): the mathematics lessons were conducted without the use of the new strategies except for the items “teacher asks students to write down the reasons for their mathematics answers”; “teacher asks students to think about the reason for solving their mathematics problems”; “teacher asks students to have more than one correct answer” and “teacher encourages students to solve mathematics questions in different ways”

The t-test of the Experimental Express with the Comparison Express, Experimental Normal Academic with the Comparison Normal Academic and Experimental Express with the Experimental Normal Academic all show no significant difference in the 95% confidence interval except the item “teacher asks students to write down their feeling about mathematics” which shows significant difference in the Experimental Express and the Comparison Express.

Post-survey results for School A. For the Comparison Express class in School A-CH(A): the mathematics lessons were conducted without the use of the new strategies except for the item “teacher asks students to have more than one correct answer”.

For the Comparison Normal Academic class in School A-CL(A): the mathematics lessons were conducted with the minimum use of new strategies except for the items “teacher asks students to write down their feelings about mathematics” ; “ teacher asks students to solve mathematics questions by themselves” and “teacher asks students to give more than one correct answer”.

When t-test was conducted to find out whether there is a change in the teaching of mathematics in the Comparison classes in School A, it was found that there is no significant difference between the way the mathematics lesson was conducted for the express class under the 95% confidence interval. However, the comparison class in the Normal Academic stream shows significant difference in two items. The two items are “teacher asks students to write down the reason for the mathematics answers” and “teacher asks students to do more than one correct answer”. Due to the reason that there is a change in subject teacher teaching the comparison class, it was assumed that the students were reflecting on the teaching styles of two different teachers. There is no clear indication that the new strategies were used in the comparison class.

Pre-survey results for School B. For the Experimental Express class in School B-EH(B): the mathematics lesson were conducted without much use of the new strategies except the item "teachers asking students to write down the reasons for their mathematics answers", "encourage students to solve mathematics in different ways" and "asking students to think about the reason for solving their mathematics problems".

For the Comparison Express class in School B-CH(B): the mathematics lessons were conducted without the use of the new strategies except "teachers encourage students to solve mathematics in different ways."

For the Experimental Normal Academic class in School B-EL(B): the mathematics lessons were conducted without the use of the new strategies except "teachers sometimes (2 to 3 times a week) encourage students to have more than one correct answer."

For the Comparison Normal Academic class in School B-CL(B): the mathematics lesson were conducted with some use of new strategies. Methods like writing down reasons; explaining in writing; using different ways to solve; thinking about reasons and to having more than one correct answer were "moderately frequent" (32%~38% of the lesson times).

The t-test of the Experimental Express with the Comparison Express, Experimental Normal Academic with the Comparison Normal Academic and Experimental Express with the Experimental Normal Academic all show no significant difference in the 95% confidence interval.

Post-survey result for School B. For the Comparison Express class in School B-CH(B): the mathematics lessons were conducted without the use of the new strategies except "teachers asking students to explain mathematics to the whole class" and "teachers encourage students to have more than one correct answer."

For the Comparison Normal Academic class in School B-CL(B): the mathematics lessons were conducted with the minimum use of new strategies. However, teachers asking students to explain mathematics to the whole class and teachers encouraging students to solve mathematics questions in different ways are considered one of the "moderately frequent" methods (35%~38% of the lesson times).

When t-test was done to find out whether there is a change in the ways of mathematics lesson in the Experimental Express class in School B, it was found that only three items are significantly different under the 95% confidence interval. The three items are "teacher asks students to explain mathematics to the class"; "Teacher asks students to write down their feelings about mathematics" and "teacher encourages students to solve mathematics in different ways".

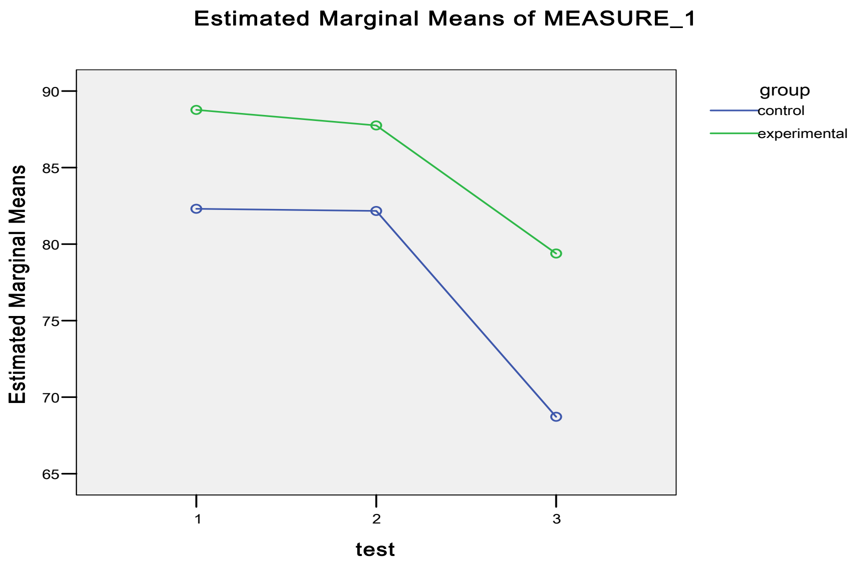
The t-test of the difference between the Experimental Normal Academic and Comparison Normal Academic shows that "teacher asks students to write down their feelings about mathematics" is the only item that shows significant difference in the 95% confidence interval.

10.4.3 School-based examination results

First, the school-based mathematics examination results were used to compare within classes and between classes for each stream in both the schools. Second, the Pre- and

Post- survey results were used to compare between classes and within classes. The examination results of the two schools are presented below.

School examination results for School A
We first start with School A as shown in Figure 10.1.



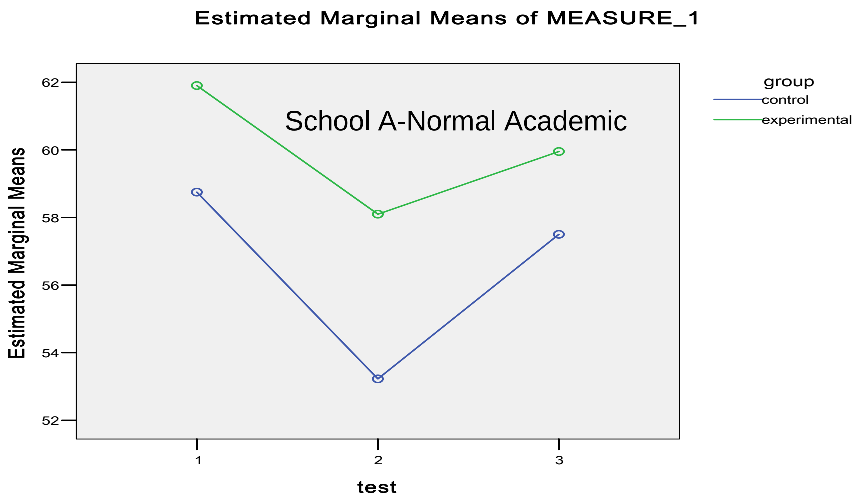


Figure 10.1. Students' results in School A in school-based tests

Note, as used in other chapters, Test 1 in the figure refers to the mid-year examination results of 2004 [M2004]; similarly Test 2 refers to the end-of-year final exam results of 2004 [F2004]; and Test 3 refers to the mid-year exam results of 2005 [M2005].

The following results were obtained about Express stream students.

1. Regarding students' PSLE math grades, students in the experimental class for Express stream had significantly higher grades than those in comparison class for Express stream.

2. In M2004, t-test results show that students in the experimental class for Express stream had significantly higher scores than those in comparison class for Express stream; In F2004, students in experimental class for Express stream had significantly higher scores than those in comparison class for Express stream; both classes showed similar trend (M2004 → F2004: drop in scores) and in terms of the extent of changes in scores, no significant difference between the two classes was detected using GLM;

3. In M2005, the significant test revealed that students in the experimental class for Express stream still had significantly higher scores than those in comparison class for Express stream; both classes had similar trend (F2004 → M2005: drop in scores) and the extent of changes in scores in experimental class for Express stream was considerably smaller than that in comparison class for Express stream using [GLM].

Now let us turn to the Normal Academic Stream.

1. Regarding students' PSLE math grades, students in the experimental class for Normal Academic stream had higher grades than those in comparison class for Normal Academic stream but no significant difference was detected;

2. In M2004, t-test results show that students in the experimental class for Normal Academic stream had higher scores than those in comparison class for Normal Academic stream but no significant difference was detected; In F2004, students in experimental class for Normal Academic stream had higher scores than those in comparison class for Normal Academic stream but no significant difference was detected; both classes showed similar trend (M2004 → F2004: a drop in scores) and in terms of the extent of changes in scores, no significant difference between the two classes was detected using [GLM];

3. In M2005, t-test shows that students in the experimental class for Normal Academic stream had higher scores than those in comparison class for Normal Academic stream but no significant difference was detected; both classes showed similar trend (F2004 → M2005: an increase in scores) and in terms of the extent of changes in scores, no significant difference between the two classes was detected using [GLM].

Comparing the experimental classes with comparison classes, it was revealed that the experimental high class is significantly more positive than comparison high class in the pre-survey and more positive in post-survey. The experimental Low class is more positive than comparison Low class in the pre-survey but is more negative in the post-survey, but no significant difference is detected.

School Examination results for School B
Figures 10.2(a) and 10.2(b) below depict the changes of students' school-based examination results in School B. From the figures, again, we shall begin with the experimental class in the Express Stream and its comparison with the comparison class.

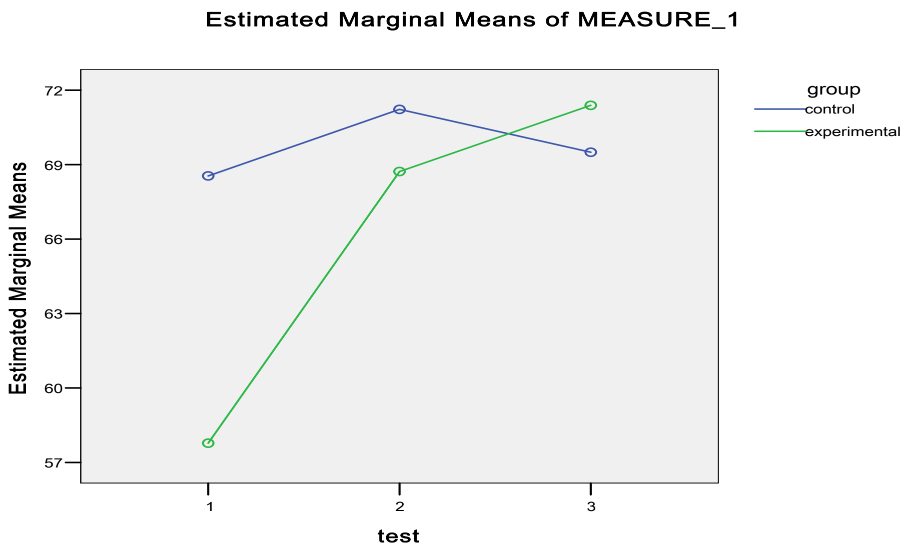


Figure 10.2(a). Students' results in School B (express) in school-based tests

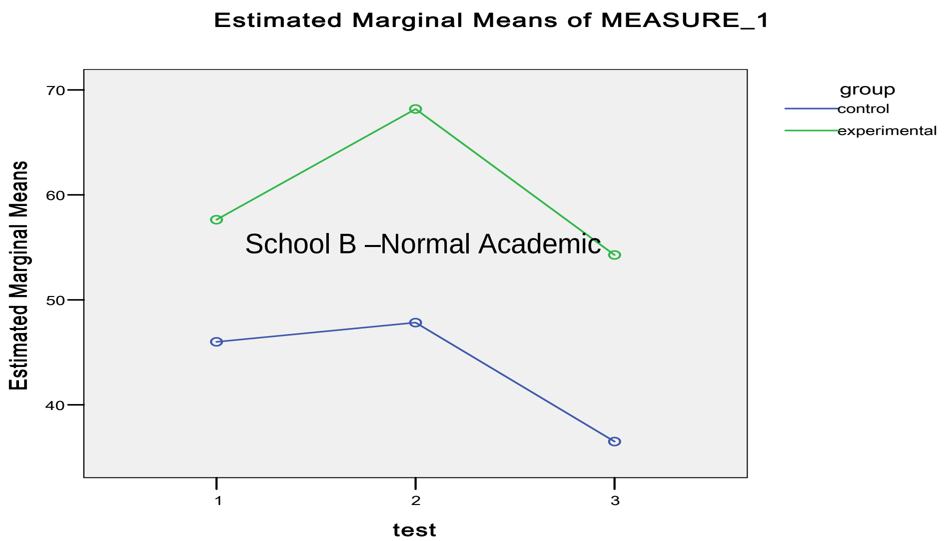


Figure 10.2(b). Students' results in School B (normal academic) in school-based tests

1. Regarding students' PSLE math grades, students in experimental class for Express stream had slightly lower grades than those in comparison class for Express stream but no significant difference was detected.

2. In M2004, the t-test results show that students in experimental class for Express stream had significantly lower scores than those in comparison class for Express stream; In F2004, students in experimental class for Express stream had lower scores than those in comparison class for Express stream; both classes showed similar trend (M2004 → F2004: an increase in scores) but the extent of changes in scores in experimental class for Express stream was significantly larger than that in comparison class for Express stream using GLM.

3. In M2005, the results show that students in experimental class for Express stream had higher scores than those in comparison class for Express stream; students in experimental class for Express stream had an increase in scores (F2004 → M2005), while those in comparison class for Express stream had a decrease in scores; in terms of the extent of changes in scores, no significant difference between the two classes was detected using GLM.

As we explained earlier, during to the changes in participating students and students, we have to study the results of the interventions into two main periods. One of the periods is from Jan 2004 to Dec 2004 for the classes 1CH(B) and 1EH(B) with 10 interventions. The other period is from Jan 2005 to May 2005 for the classes 2CH(B) and 2EH(B) with another 10 interventions. The two periods had different experimental groups and comparison groups due to the streaming exercise in the school. Because of the changes in 2005, it was thought that, the data after year 2004 will not be valid and needed to be discarded. However, after examining the status of the classes in terms of students' profile and teachers' teaching experience, we found that the experimental class and the comparison class in the Express stream were still very similar in 2005. The results can be meaningful to us due to the significant improvement in school examination results for the experimental classes. There is a significant interaction between the school examination results in these two classes. 2CH(B) was better than 2EH(B) at the start of 2005. However, the mid-year school examination result showed that 2EH(B) performed better than 2CH(B). We are not sure whether there was Hawthorne effect or Pygmalion effect in this case. However, there are altogether 11 interventions in 2004 and 11 interventions in 2005, if we consider that the interventions were done separately for two different groups of samples where the two samples had an intersection of 18 students, there are outcomes that may worth spending some time considering the effect that interventions might have with the students.

Now let us look at the Normal Academic Stream.

1. Regarding students' PSLE math grades, students in experimental class for Normal Academic stream had lower grades than those in comparison class for Normal Academic stream but no significant difference was detected.

2. In M2004, the t-test results show that students in experimental class for Normal Academic stream had lower scores than those in comparison class for Normal Academic stream but no significant difference was detected; In F2004, students in experimental class for Normal Academic stream had lower scores than those in comparison class for Normal Academic stream but no significant difference was detected; both classes showed similar trend (M2004 → F2004: increase in scores) but the extent of changes in scores in experimental class for Express stream was significantly larger than those in comparison class for Normal Academic stream using GLM.

3. In M2005, the significant test shows that students in experimental class for Normal Academic stream had significantly higher scores than those in comparison class for Normal Academic stream ; both classes showed similar trend (F2004 → M2005: drop in scores) and in terms of the extent of changes in scores, no significant difference between the two classes was detected using GLM.

By comparing the experimental class and comparison class, we can find that the experimental High class is more positive than comparison High class in both the Pre- and Post-survey. The experimental Low class is more negative than comparison Low class in both the Pre- and Post-survey. Therefore, it is difficult to make a conclusive judgment, but again it seems safe to say that the intervention did not produce negative influences on students learning based on standard school-based examinations.

10.4.4 Questionnaire on self–assessment strategy

The survey results on the students' attitude towards the new strategy show that most of the students in the Express classes in both schools are 'neutral' or 'agree a little' to use the new strategies in their mathematics lessons. However, all the students in the Normal Academic classes in both schools stay 'neutral' to the use of the new strategies in their mathematics lessons, which seems that self-assessment can be better accepted for high-achieving students.

Below is a summary of the main findings about students' views on the self-assessment strategies in each participating class.

EH(A): Most of the students responded that they are 'neutral' or 'agree a little' to the use of the new strategies in the mathematics lesson.

EL(A): All the students responded that they are 'neutral' to the use of the new strategies in their mathematics lesson.

EH(B): Most of the students responded that they are 'neutral' to the use of the new strategies in their mathematics lesson.

EL(B) : All the students responded that they are 'neutral' to the use of the new strategies in their mathematics lesson.

Interestingly, the questionnaire survey on students' views about the self-constructed test activity found that students enjoyed it, to a large degree. Below is a summary of the survey results (See the items in Appendix 10.4)

Question 1

(a) Did you group submit the question?

	Yes	**No**
Number of students	38	0

(b) If the answer to (a) is "yes", did the question your group submitted appear in the test?

	Yes	**No**
Number of students	16	22

(c) If the answer to (b) is "yes", the question appeared in the test as test item ______ (Please write down the item number if you can still remember).

The number of students who gave the item number is denoted in brackets:

a. Item number '3'. (3)
b. Item number '4'. (1)
c. Item number '5'. (2)
d. Item number '6'. (4)
e. Blank. (28)

When you found the item you designed in the test, what did you feel? (check ALL you think true)

Items	Number of students
I felt good about it	5
I felt excited about it	6
I felt proud of it	6

21 out of the 22 students who answered 'No' in part (b) left this portion blank. There is one student, though he/she answered 'No' in part (b), put a tick against the item 'I felt excited about it'.

Other feelings:
a. 'Sad as I forgot the answer.'
b. 'Nervous.'
c. 'But I didn't really know how to do it.'
d. 'I was going "Oh My God?!!"'
e. 'The question is not submitted by me. We split ourselves into two smaller groups and the question is submitted by my friend. I am not sure about how they do it.'
f. 'Surprised.'
g. 'No feelings actually, the question was quite easy.'

Question 2
Please estimate how many hours you spent in designing the question in total (in your group and individually)? _____ hours

The number of students who gave the number of hours is denoted in brackets
a. 1/6 hours. (1)
b. ¼ hours. (1)
c. ½ hours. (2)
d. Less than ½ hour. (1)
e. 1 hour. (23)
f. 2 hours. (8)
g. 3 hours. (1)
h. Blank. (1)

Question 3
How did your group design the questions? (check ALL you think true)

Items	Number of students
a. We reviewed the topic covered first, then had group discussion.	25
b. We asked one group member to design it. Other actually did not contribute.	1
c. We just copied down a question from a book.	2
d. We went through a lot of similar questions in the available resources, then we selected a question from there and modified it to be a new question.	11
e. We did not know how to start, so we asked the teacher for help.	2
f. We created the question together.	22

Others:
a. We created the questions individually.
b. Each group member designed a question and we handed up all the questions.
c. We look through the topic and design the question.
d. We divided ourselves into two subgroups and did one question each. In my subgroup, one of us created the diagram and the other two created the answers.
e. Then we each come out with two questions and pass up to the teacher in a group.

Question 4
What did you gain and/or learn from designing the question? (check ALL you think true)

Items	Number of students
a. I reviewed the topic relevant to the question.	21
b. I know better how to solve mathematical questions in this topic.	20
c. I know better about my classmates.	11
d. I learned how to work together with others	23

e. I did not gain/learn anything. It was a waste of time.	0

Others:
a. I learnt that my classmates could set very hard questions.
b. Creating maths problems can be fun!

Question 5
(a) Was it the first time that you created mathematical question by yourselves?

	Yes	No
Number of students	16	22

(b) Would you like to do it again?

	Yes	No
Number of students	22	15

1 student put 'Maybe'.

(c) If your answer to (b) is "yes", then how frequently should we have it in our class?

Items	Number of students
a. Once a week.	3
b. Once a month.	2
c. Once a term.	8
d. Once a semester.	2
e. Once for each chapter.	9

14 out of the 15 students who answered 'No' in part (b) left this portion blank. There is 1 student, though he/she answered 'No' in part (b), put a tick against the item 'Once a term'.

Others:
a. Once every two chapters.

Question 6
Overall, how much do you think it was beneficial for you to design the question?

Items	Number of students
a. Very much.	7
b. Somehow.	24
c. A little bit.	6
d. Not at all.	0

1 student put 'Much'.

Question 7
How did you feel about the test? (check ALL you think true)

Items	Number of students
a. It was too difficult.	14
b. It was too easy.	0
c. It was somehow interesting to solve test items set by other classmates.	24

Other feelings:
a. Some too difficult, some alright but just no easy.
b. A little too complicated and fair to those who set the question but unfair to others.

Question 8

(a) Was it the first time you took a test mainly constructed by yourselves or your classmates?

	Yes	No
Number of students	32	6

(b) Would you like to do it again?

	Yes	No
Number of students	18	19

1 student put 'Maybe'.

(c) If your answer to (b) is "yes", then how frequently should we have it in our class?

Items	Number of students
a. Once a week.	1
b. Once a month.	2
c. Once a term.	5
d. Once a semester.	6
e. Once for each chapter.	5

Others:

a. Once for two chapters.

Question 9

If you wish to tell me anything else about the activities of designing the questions and the test, please feel free to write it down.

Summary: 33 students left this question blank or put 'Nil'. 2 students felt that the questions were too difficult. 1 student commented that it would be more interesting if every student get to do every question they designed. Another student felt that the questions set by the teacher are more fun and the questions set by the students are a bit too challenging.

a. Blank or 'Nothing'. (33)
b. 'I think that it would be more interesting if we pass our questions around the class as a quiz, in that manner, everyone's questions would be used'. (1)
c. 'I feel that questions set by teachers are finer as the questions by students are a little too challenging'. (1)
d. 'Maybe if we have time, we can find some interesting maths problems to play with'. (1)
e. 'Some questions are too difficult'. (2)

In general, we think that engaging students in self-constructed test can be a useful way for assessment as well as learning, but it must be used wisely, including making the requirement clear, giving adequate instructions and help, using and editing students' questions properly, and not using it too frequently.

10.4.5 Self-assessment tests

Attitudes towards Self-assessment

As said earlier, the results were obtained from students' responses in the experimental classes to Part B in post-survey, containing 21 items.

We present the results in each school respectively.

School A. Students on the whole provided positive responses to the majority of the items in this part; in particular, it was the case for 19 items (including 1 for Neutral) for experimental/high class and 15 items for experimental/low class.

Students in experimental/high class provided more negative responses than those in experimental/low class on all but 9 items (Item 24, Item 25, Item 26, Item 30, Item 32, Item 38, Item 39, Item 40, and Item 42); moreover, experimental/high class provided significantly more positive responses than their counterparts on 1 item (Item 39).

School B. Students on the whole provided positive responses to the majority of the items in this part; in particular, it was the case for 14 items for experimental class for Express stream and 12 items for experimental class for Normal Academic stream.

Students in the experimental class for Express stream provided more positive responses than those in experimental class for Normal Academic stream on all but 6 items (Item 23, Item 24, Item 25, Item 32, Item 34, Item 35) but no significant difference was detected.

Pre- and Post- new strategy tests
We noted that the performance of the post–test was not as good as the pre-test in the new strategy tests in general. One reason for this might be the fact that, for the post-test, students already knew that their performance in the test did not really count in their school report, which could affect their attitude and behavior in taking the test.

Nevertheless, the experimental classes were consistently better than the comparison classes in both the pre- and post- test for School A and the experimental classes are slightly weaker than the comparison classes in School B. However, all the differences do not give significant difference in their performance.

10.4.6 Interviews

Regarding the findings from interviews, we start with interviews with teachers, followed by interviews with students.

Interviews with teachers
The interviews show that all the teachers involved in the study are very supportive of the self-assessment strategy.

One of the teacher of the experimental classes also used the self-assessment strategy in her other classes with another teaching subject. She said she had experienced the benefit of using the self-assessment strategy in her teaching. She claimed that “It was a routine for my average ability class to fill in the self-assessment forms after every chapter had been taught. They were rather uncomfortable initially but they began to appreciate the rationale after explaining to them. This served as corrective feedback and it would be most useful if it was immediate, frequent and communicated in non-judgmental ways.”

The interview revealed that the teachers in School A have decided to implement the self-assessment to all their lower secondary students. They are confident that it is going to benefit their students even through the research result has not be published. In School B, teachers involved in the study also believe that the self-assessment is very useful. It instills in students the learning responsibility and establishes a communication channel for teacher to teacher better and students to learn better. It benefited both the teachers and the students in general.

Interviews with students
During the experimental period, teachers in the experimental class use evaluation and reflection worksheets to find out their students’ learning difficulties and use the findings to improve on their teaching. As mentioned earlier, students were asked to construct questions on the topics they learnt. Some of their questions were chosen or were amended to form part of their test papers. Students were very excited when their questions were chosen for their class test.

Below are excerpts from interviews:

> Teacher A: How did you feel when you were preparing these questions?
> Student E: Exciting.
> Student F: Very exciting.
> ……………………………………………….
>
> Teacher B: And I don't know whether you remember what the feeling was like when I mentioned to you all that you are preparing the test questions for yourself?
> Student F: Hah? I was very shocked.
> ………………………………………………
>
> Teacher B: Do you like this idea... for another test?
> Student E: Of course.

Researchers also interviewed three students per experimental class. Results of the interviews show that students generally treasured the opportunities given to improve their learning and to clarify many of their doubt. They also gave suggestions to fine tune the ways self-assessment can be carried out in their classed. Generally, all of the students interviewed supported the idea of implementing the self-assessment strategy in their classroom.

10.5 Conclusions, Implications, and Recommendations

The sub-study revealed that the teachers in the experimental class were very confident in their classroom teaching and were very natural in the integrating of the self-assessment into their classroom teaching. Their students reacted very normally even when the researchers were video recording their lesson.

All the teachers who were involved in the intervention (in both schools) held very positive views that self-assessment is a good strategy that would benefit in the teaching and learning of mathematics. Teachers concerned were confident that the strategy can be integrated into their mathematics classrooms. Generally, both teachers and students felt the self-reflection can be done regularly as a routine activity. Through regular reflection, students can learn to be reflective in their learning and got the chance to clarify the doubts and concepts. Teachers improve their teaching by responding to students' reflection. When students do their self-evaluation per chapter or per regular interval, it helped them become better reflective learners.

With the help of new assessment strategies, both teachers and students became more critical and focused in their teaching and learning. Through self-evaluation, students also learned to correct their own mistakes and ways to improve their own learning. With the self-constructed test, students were able to revise their work spontaneously and set their own learning target as well as challenging their own learning.

In short, we would say that effectively using engaging students in self-assessment strategies in the mathematics classrooms can not only help teachers understand better students' learning and their own teaching, but also provide students with the meaningful opportunities to reflect on their own learning, and hence improve teachers' teaching and student' learning. It is our belief after this sub-study that student self-assessment can and should be done as an integral or routine activity in the teaching and learning of mathematics at the secondary levels.

Chapter 11 Summary, Conclusions, and Recommendations

In this final chapter, we shall first provide a summary of this project, then draw some overall conclusions based on what we have presented in the previous chapters, and finally, offer relevant suggestions and recommendations for policy makers, school administrators, mathematics teachers, and researchers concerning the use of new assessment strategies in mathematics classrooms.

11.1 Summary and Conclusions

The MAP project aimed to provide research-based evidence and practical suggestions for promoting the effective use of new assessment strategies in Singapore mathematics classrooms. There were three broad research questions for the MAP project:

1. What are the influences of “new assessment strategies” on students’ learning of mathematics in their cognitive domain?

2. What are the influences of “new assessment strategies” on students’ learning of mathematics in their affective domain?

3. How can “new assessment strategies” be effectively integrated into mathematics classrooms?

In this study, largely consistent with NCTM’s definition (NCTM, 1995), we view assessment as the process of teachers’ gathering information about a student’s knowledge of, ability to use, and disposition toward mathematics, mathematics teaching and mathematics learning, and of their making inferences from the information gathered for or abut students’ learning in mathematics. In connection to classroom instruction, assessment is therefore defined as the process of teachers’ gathering information about students’ learning, which include their achievement and behavior in both cognitive and affective domains, and hence making informed decisions for classroom instruction. For classroom teachers, the ultimate purpose of assessment is to improve the quality of teaching and learning. In relation to this, we view assessment as an essential part of the process of teaching and learning, and it should be integrated into classroom instructional activity.

Concerning alternative assessment strategies or, as we called in this study, new assessment strategies, there has been no universally agreed definition, nor has there been a universal classification, except that people generally agree that alternative assessment is different from traditional paper-and-pencil tests.

In the MAP project, we used the term alternative assessment in mathematics to include the following specific techniques in student assessment: performance assessment, authentic assessment, portfolio assessment, journal writing assessment, project assessment, oral presentation assessment, interview assessment, classroom observation assessment, student self-assessment, student-constructed assessment, among others.

In particular, The MAP project focuses on four alternative assessment strategies: project assessment, performance assessment, student self-assessment, and communication assessment.

Performance assessment, or performance-based assessment, refers to the assessment practice in which the information about students’ learning is gathered through students’ work on performance tasks. Performance tasks in this study mainly include authentic real-life problems and open-ended tasks. Project assessment, or project-based assessment, refers to the assessment practice in which the teacher gathers the information about students learning through their work on project tasks. Student self-assessment refers to the assessment practice in which the information about students' learning is gathered through their reflection, evaluation, and report to the teacher. Communication assessment in this study refers to the assessment practice in which the information about students' learning is gathered through students' performance on communication tasks, including mainly both journal writing (written communication) and oral presentation (oral communication) tasks.

The MAP project involved classroom-based intervention over a period of about three school semesters in 8 primary and 8 secondary schools, including both high-performing schools and typical neighborhood schools. To have a benchmark to better detect the influences of the use of new assessment strategies in the participating classes (or the so-called "experimental classes" for convenience) which received intervention, whenever possible comparison classes were selected in the same school, same stream, and same grade, using same textbooks, etc. Nevertheless, one should note that the comparison classes were not really controlled, or even expected, as "control groups" like in a traditional lab-based experimental study, and in particular, it was impossible in many cases to have equivalent groups with regard to the professional and/or academic background of teachers and students. This situation is understandable, and in a large sense, to us, more realistic and meaningful, given the authentic context, the focus, scale, and duration of the MAP project.

A number of instruments were designed for the project. The data were mainly collected through pre- and post- questionnaire surveys, pre- and post- tests, classroom observations, interviews with teachers and students, and school-based examination scores.

While different sub-study of the MAP project revealed different conclusions for each specific assessment strategy it addressed, we think overall there are three main conclusions that can be drawn from the MAP project:

First, the qualitative data (interviews, observations, video-recordings) endorse the view that new assessment tasks appear helpful in developing students' higher-order thinking, communication skills, self-regulation, and self-reflection in learning. The most encouraging part of the evidence came from two facts: (1) given adequate time and help, both teachers and students were capable of working on the new assessment strategies, and (2) both teachers and students offered overall very positive views about the value and feasibility of integrating these new assessment strategies into their daily teaching and learning activities.

Second, the quantitative data including survey questionnaire, pre- and post- test results, etc., indicate that most intervention classes performed better or equally good as comparison classes in both cognitive and affective domains.

Third, overall, while there is no conclusive finding about the positive impact of the intervention on student exam-based performance because of the complexity of the factors (e.g., teachers' experience and skills) that might affect students' academic achievements, the research team believe that it is safe to conclude that the use of the new assessment strategies will not adversely affect students' academic achievement based on regular school exams.

Table 11.1 provides a summary of the change of the results on school-based exams in primary and secondary schools respectively, from the mid-year of 2004 (M2004) to the mid year of 2005 (M2005).

Table 11.1. A summary of results on school-based exams

	Primary							
	M2004 → F2004				F2004 → M2005			
	CO	PT	PW	SS	CO	PT	PW	SS
+ (sig)							X	
+	X X	X		X X		X	X	
-	X X	X X	X		X X X	X X		X X
- (sig)			X		X			
	Secondary							
	M2004 → F2004				F2004 → M2005			
	CO	PT	PW	SS	CO	PT	PW	SS
+ (sig)	X	XX		XX	XX	XX		X
+	X	XX	X	X	X	X	X	X
-	X		XX	X	X	X	X	XX
- (sig)							X	

Note. In the table, all the changes are taken with corresponding comparison classes as a base-line. Each "X" represents one case (class). "+" means that the change is in favor of the experimental class but is not statistically significant based on t-test, while "-" means that the change is in favor of the comparison class without statistical significance. "+ (sig.)" and "- (sig.)" denote statistical significance of the difference.

The third conclusion above is particularly true, if one takes into account the fact that new assessment tasks were not included in the participating schools' regular examination practice for students' final grades. As reported in the previous chapters, this fact posed a challenging environment and, to some degree, has affected both teachers' and students' implementing new assessment tasks, and hence produced negative influences on the impact of the invention. Nevertheless, to reach a more specific conclusion in this aspect, we think a further study with a different focus, research design, and a smaller-scale is needed.

11.2 Suggestions and Recommendations

According to what we observed from the MAP project, we believe that the effectiveness of the use of new assessment strategies in the mathematics learning should not be tied to the performance in the school mathematics achievement tests based on the traditional written tests. They are other important advantages of using the new assessment in mathematics learning that are often not reflected in the regular and traditional achievement tests.

Educators, especially these who wish to reform classroom assessment practice, should realize that many factors bear on students' academic performance. These factors include, for example, student academic background, peer influence, family socio-economic status, school and class environment, teachers, textbooks, etc. Hence, the value of new assessment strategies should be reasonably understood. In a large sense, new assessment strategies are more for better teaching and learning, not for traditional testing.

The MAP study also revealed that both teachers and students had challenging time for getting started. It took a period of time, usually from a few weeks to a few months, for them to get familiar with the concept, value, methods, and skills about new assessment strategies. For example, students especially those in the low ability class were not confident in doing mathematics project and communication work.

Therefore, for school teachers especially for those who are relatively new to these assessment strategies, we suggest they start the new assessment practice with small steps, and move on gradually (for example, giving guided and less challenging project tasks at the beginning), so the effort can be sustained and achieve more reachable results (Remember it is much easier to go back to old practice).

In the process of implementing new assessment tasks, we believe it is particularly important for teachers to make clear to students what is expected for both the process and end-product at the beginning, give enough guidance and help during students' work, and offer timely feedback in the final stage, so the assessment can be more effective. Students' possible difficulty and confusion in working on new assessment tasks should not be underestimated.

Meanwhile, professional development must be provided to better prepare teachers for the use of new assessment strategies in the mathematics classrooms. The MAP project offered three project-wide workshops for all participating schools and found they were very helpful and highly welcomed by the participating teachers, HODs, and principals. Besides in-service training offered by the National Institute of Education or relevant units of the Ministry of Education, we think there is a need to have more teachers' own exchange and sharing in this direction.

For the school administrators, we think their support for teachers to try new assessment strategies is critical. In particular, it is important to create a school environment and culture that appreciates the value of new assessment strategies, and understands that the effects of using new assessment tasks on student learning might not be readily reflected in the regular semestral examinations. We are glad to know that more and more schools have recently started using new assessment tasks as part of students' continual assessment (CA), and we support such effort.

In addition, relevant resources must be made accessible to teachers. The MAP project found that one of the most challenging tasks for the researchers was to design the intervention tasks and assessment criteria or rubrics for these tasks. We noted that there have been an increasing number of online resources in new assessment strategies available for school teachers to use, however, most of them are from other countries, developed based on their own curriculum and educational environments.

Although over the last few years, there has also been an increase in the resources related to new assessment strategies in Singapore and, in particular, some new school textbooks have integrated some new assessment tasks (e. g., see Fan & Zhu 2007), we think the resources developed for local teachers and students are still far from enough, which pose a real challenge for schools and teachers to implement the new assessment strategies.

Finally and we believe also most importantly, new assessment strategies should be integrated into the teaching and learning process. In particular, the tasks should 'replace' some other traditional and less important tasks and be an integral part of, but not 'add on' to, the mathematics curriculum and instruction. To achieve this goal, policy makers, curriculum developers, school administrators, and classroom teachers need to share a common vision about the value of education as well as assessment and hence make concerted and sustained effort.

Acknowledgements

This paper makes use of data from the research project "Integrating New Assessment Strategies into Mathematics Classrooms" (CRP 24/03 FLH), funded by the Centre for Research in Pedagogy and Practice, National Institute of Education, Singapore. The views expressed in this paper are the authors' and do not necessarily represent the views of the Centre or the Institute.

It is impossible to undertake a project like the Mathematics Assessment Project (MAP) without the support of many people. The authors would like here particularly to acknowledge three groups: the school teachers and administrators, the Centre for Research in Pedagogy and Practice (CRPP) of the National Institute of Education (NIE), and our fellow team members.

First, the authors express their appreciation to all the school teachers and administrators who participated or helped us in this research project (including those who helped for piloting research instruments), though because of the agreement for conducting this study we cannot explicitly their names here. In particular, the teachers had the most difficult task. Not only did they have to satisfy the needs of the project as outlined by the contact person from our research team but also they had to deliver a programme that met the needs of their students and the school administration. They had to keep the project going and at the same time ensure that their students covered the curriculum and did well in the various school assessments. Without teachers and school administrators' commitment and support, a school-based project like the MAP cannot succeed.

Second, the authors are very grateful to the CRPP and its management team and administrative staff for all the support it has offered to the MAP. For this project, CRPP provided not only the critical financial support for the project, but also gave us all the professional trust and operational help we needed. CRPP's financial, administrative, and moral support for this project is deeply appreciated.

Finally, the authors must thank their fellow team members. The project with planning, implementation and continuous contact with the schools spread over more than two years (including having regular team meetings almost every month). To commit oneself to any project for this period takes both an extreme effort as well as a belief in the project. Without their continual interaction and support it would not have been possible to bring such a large project to a successful conclusion.

Notes

1 "Students" and "Pupils" are used interchangeably in this report.

2 In Singapore schools, the Scheme of Work, usually developed by the department of mathematics, is a school-based document that spells out the syllabus to be covered for each term.

3 This sub-project is under the charge of Lionel Pereira-Mendoza.

4 Please note that although Chapter 3 to Chapter 10 each report a different sub-study, they contain some commonalities or overlapping in explanation about the general shared conceptions and research methods. In a sense, they are kept with the purpose to maintain the integrity and self-completeness of the report for each sub-study for easy reading and reference.

5 This sub-project is under the charge of Quek Khiok Seng.

6 This sub-project is under the charge of Koay Phong Lee.

7 This sub-project is under the charge of Juliana Donna Ng Chye Huat.

8 This sub-project was under the charge of Zhu Yan.

9 Readers who are interested may contact the researchers for the detailed frequency tables of students' responses on all the questionnaire items for both the pre- and post-surveys.

10 Effect size r is calculated when significant difference is detected. According to Cohen (1992, 1988), an r value over .5 is considered to be "large", around .1 to be "weak", and around .3 to be "medium".

11 This sub-project is under the charge of Tan-Foo Kum Fong. A more detailed report can be found in Foo, K. F. (2007). *Integrating Performance Tasks in the Secondary Mathematics Classroom: An Empirical Study*. Unpublished Master's Dissertation, National Institute of Education, Nanyang Technological University, Singapore.

[12] The experimental (E) and comparison (C) class were selected by the school administration for this study. The researcher had informed the school administration that she would work with any class and the mathematics teacher of the class before the study commenced.

[13] The feature of shaded rows denotes negatively worded items and is the same for all the corresponding tables below.

[14] The generic rubric was adapted from the site, http://rubistar.4teachers.org/index.php

[15] This sub-project is under the charge of Yeo Shu Mei. A more detailed report can be found in Yeo, S. M. (2008). *Using Journal Writing and Oral Presentation as Alternative Assessment Strategies in Mathematics Instruction: An Empirical Study in Singapore Secondary Schools*. Unpublished PhD Thesis, National Institute of Education, Nanyang Technological University, Singapore.

[16] This sub-project is under the charge of Fan Lianghuo and Teo Soh Wah.

References

Adam, T. L. (1998). Alternative assessment in elementary school mathematics. *Childhood Education, 74*(4), 220–224.

Badger, E. (1992). More than testing. *Arithmetic Teacher, 39*(9), 7–11.

Bagley, T., & Gallenberger, C. (1992). Assessing students' dispositions: Using journals to improve students' performance. *The Mathematics Teacher, 85*(8), 660–663.

Bayles, E. E. (1950). *The theory and practice of teaching.* New York: Harper & Brothers.

Brown, J., & Shavelson, R. (1994). Does your testing match your teaching style? *Instructor, 103*(7), 86–89.

Buechler, M. (1992). *Performance assessment.* Policy Bulletin, No. PB-B13.

Clarke, D. J. (1997). *Constructive assessment in mathematics: Practical steps for classroom teachers.* Berkeley, CA: Key Curriculum Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Dougherty, B. J. (1996). The write way: A look at journal writing in first-year algebra. *Mathematics Teacher, 89*(7), 556–560.

Drake, B., & Amspaugh, L. (1994). What writing reveals in mathematics. *Focus on Learning Problems in Mathematics, 16*(3), 43–50.

Fan, L. (2002). In-service training in alternative assessment with Singapore mathematics teachers. *The Mathematics Educator*, *6*(2), 77–94.

Fan L. (2005). Improving mathematics teaching and learning through effective classroom assessment: Experiences and perspectives from Singapore schools. In N. Y. Wong (Ed.), *Revisiting Mathematics Education in Hong Kong for the New Millennium* (pp. 359–376). Hong Kong: Hong Kong Association for Mathematics Education.

Fan, L., & Yeo, S. M. (2000). Journal writing as an alternative assessment. *Scientas (Journal of the Science Teacher Association of Singapore), 34,* 16–25.

Fan, L. (2006). Making alternative assessment an integral part of instructional practice. In P. Y. Lee (Ed.), *Teaching secondary school mathematics*: *A resource book* (pp. 343–354). Singapore: McGraw Hill.

Fan, L., Quek, K. S., Ng, J. C. H, et al. (2006). *New assessment strategies in mathematics: An annotated bibliography of alternative assessment in mathematics*. Singapore: Centre for Research in Pedagogy and Practice (CRPP), National Institute of Education; 112 pages. (Note: This work is part of MAP project with contribution from 15 team members; ISBN: 981-05-5-61-8)

Fan, L., & Zhu, Y. (2000). Problem solving in Singaporean secondary mathematics textbooks. *The Mathematics Educator*, *5*(1/2), 117–141.

Fan, L., & Zhu. Y. (2007). From convergence to divergence: The development of mathematical problem solving in research, curriculum, and classroom practice in

Singapore. *ZDM-The International Journal on Mathematics Education*, *39* (5&6), 491–501.

Foo, K. F. (2007): *Integrating Performance Tasks in the Secondary Mathematics Classroom: An Empirical Study*. Unpublished master's dissertation, National Institute of Education, Nanyang Technological Univerisity, Singapore.

Ginsburg, H. P., Lopez, L. S., Mukhopodhyay, S., Yamamoto, T., Willis, M., & Kelly, M. S. (1992). Assessing understandings of arithmetic. In R. Lesh & S. Lamon (Eds.), *Assessments of authentic performance in school Mathematics* (pp. 17–62). Washington: American Association for the Advancement of Sciences Press.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research. 77*(3), 334–372.

Gripps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.

Hadjioannou, X. (2007). Bringing the background to the foreground: What do classroom environments that support authentic discussions look like? *American Educational Research Journal, 44*(2), 370–399.

Haines, C., & Izard, J. (1994). Assessing mathematical communications about projects and investigations. *Educational Studies in Mathematics*. *27*(4), 373–386.

Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V. S., et al. (1997). *Performance assessment in IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Hart, L. E., & Walker, J. (1993). The role of affect in teaching and learning mathematics. In D. Owens & S. Wagner (Eds.), *Research ideas in the classroom: Middle grades mathematics* (pp. 22–38). New York: Macmillan.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33–34.

Howe, A. C., & Jones, L., (1998). *Engaging children in science* (2nd ed.). Upper Saddle River, NJ: Merrill.

Kane, M. B., Khattri, N., Reeve, A. L., & Adamson, R. J. (1997). *Assessment of student performance*. Washington, DC: Studies of Education Reform, Office of Education Research and Improvement, U.S. Department of Education.

Kulm, G. (1994). *Mathematics assessment: What works in the classroom*. San Francisco, CA: Jossey-Bass.

Lester, F. K., Garofalo, J., & Kroll, D. L. (1989). Self-confidence, interest, beliefs and metacognition: Key influences in problem solving behaviour. In D. B. Mcleod & V. M. Adams (Eds.), *Affect and mathematical problem solving* (pp. 75–88). New York: Springer-Verlag.

London, R. (1993). *A curriculum of nonroutine problems*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. ED359213).

Macnab, D. S., & Payne, F. (2003). Beliefs, attitudes and practices in mathematics teaching: Perceptions of Scottish primary school student teachers. *Journal of Education for Teaching*, *29*(1), 55–68.

Ministry of Education. (2000). *Mathematics syllabus (primary).* Singapore: Curriculum Planning and Development Division.

Ministry of Education. (2001). *Project work to be included for university admission in 2005* [Press Release 20 June]. Singapore: MOE. Retrieved 28 Sept, 2007, from http://www.moe.gov.sg/press/2001/pr20062001.htm

Ministry of Education. (2002). *Mathematics syllabus* (*lower secondary*). Singapore: Curriculum Planning and Development Division.

Ministry of Education. (2004a). *Assessment guidelines (primary mathematics).* Singapore: Testing and Assessment Branch, Research and Testing Division, MOE.

Ministry of Education. (2004b). *Assessment guidelines (lower secondary mathematics)*. Singapore: Testing and Assessment Branch, Research and Testing Division, MOE.
National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.
Ng, C. H. J. (2003). Introduction: A journal-based assessment. *Mathematics Newsletter, 6* (MOE, Singapore). Retrieved January 9, 2004, from http://intranet.moe.edu.sg/maths/newsletter/
Nisbet, J., & Shucksmith, J. (1984). *The seventh sense*. Edinburgh: SCRE Publication.
Pearce, d. L., & Davison, D. M. (1988). Teacher use of writing in the junior high mathematics classroom. *School Science and Mathematics, 88*(1), 6–15.
Pugalee, D. (2001). Writing, mathematics, and metacognition: Looking for connections through students' work in mathematical problem solving. *School Science and Mathematics, 101*(5), 236–244.
Raymond, A. M. (1994). Assessment in mathematics education: What are some of the alternatives in alternative assessment. *Contemporary Education*, *66*(1), 13–17.
Richardson, K. (1988). Assessing understanding. *Arithmetic Teacher*, *35*(6), 39–41.
Robinson, D. (1998). Student portfolios in mathematics. *Mathematics Teacher, 91*(4), 318–325.
Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook on mathematics teaching and learning* (pp. 334–370). New York: Macmillan.
Seto, C. (2002). Oral presentation as an alternative assessment in mathematics. In D. Edge & B. H. Yeap (Eds.), *Mathematics education for a knowledge-based era (Vol. 2)* (pp. 33–39). Singapore: Association of Mathematics Educators.
Shield, M., & Galbraith, P. (1998). The analysis of student expository writing in mathematics. *Educational Studies in Mathematics, 36*(1), 29–52.
Stacey, K. (1987). What to assess when assessing problem solving? *Australian Mathematics Teacher*, *43*(3), 21–24.
Stempien, M., & Borasi, R. (1985). Students' writing in mathematics: Some ideas and experiences. *For the Learning of Mathematics*, *5*(3), 14–17.
Stenmark, J. (1991). *Mathematics assessment: Myths, models, good questions, and practical suggestions.* Reston, VA: National Council of Teachers of Mathematics.
Stenmark, J. K., Beck, P., & Asturias, H. (1994). Assessment: A Room with More than One View. *Mathematics Teaching in the Middle School*, *1*(1), 44–49.
Stephens, M. (2004). Research and development in external assessments. Paper presented at the Tenth International Congress on Mathematical Education (ICME-10), Copenhagen, Denmark.
Wisconsin Education Association Council. (1996). *Performance assessment*. Education Issues Series.
Wong, N.-Y., Lam, C.-C., Wong, K.-M. P., Leung, F. K.-S., & Mok, I. A.-C. (2001). Students' views of mathematics learning: A cross-sectional survey in Hong Kong. *Education Journal*, *29*(2), 37–59.
Wu, H. (1994). The role of open-ended problems in mathematics education. *Journal of Mathematics Behavior*, *13*(1), 115–128.
Yazilah, A., & Fan, L. (2002). Exploring how to implement journal writing effectively in primary mathematics in Singapore. In D. Edge & B. H. Yeap (Eds.), *Mathematics education for a knowledge-based era* (Vol. 2) (pp. 56–62). Singapore: Association of Mathematics Educators.
Yeo, S. M. (2001). *Using journal writing as an alternative assessment in junior college mathematics classrooms.* Unpublished master's dissertation, Nanyang Technological University, Singapore.

Yeo, S. M. (2008): *Using Journal Writing and Oral Presentation as Alternative Assessment Strategies in Mathematics Instruction: An Empirical Study in Singapore Secondary Schools.* Unpublished PhD Thesis, National Institute of Education, Nanyang Technological Univerisity, Singapore.

Yeo, S. M., & Zhu, Y. (2005). *Higher-order thinking in Singapore mathematics classrooms.* Paper presented at the Educational Conference: Redesigning pedagogy in research, policy, practice. Singapore: Centre for Research in Pedagogy and Practice, National Institute of Education.

Zehavi, N., Bruckheimer, M., & Ben-Zvi, R. (1988). Effect of assignment projects on students' mathematical activity. *Journal for Research in Mathematics Education*, *19*(5), 421–438.

About the authors

Fan Liang Huo is Associate Professor with the Mathematics and Mathematics Education Academic Group at the National Institute of Education, Nanyang Technological University, Singapore.

Quek Khiok Seng is Assistant Professor with the Psychological Studies Academic Group at the National Institute of Education, Nanyang Technological University, Singapore.

Koay Phong Lee is Associate Professor with the Mathematics and Mathematics Education Academic Group at the National Institute of Education, Nanyang Technological University, Singapore.

Lionel Pereira-Mendoza is Coordinator of the Doctor of Education (EDd) Programmes with the Graduate Programmes and Research Office at the National Institute of Education, Nanyang Technological University, Singapore.

Contact us

For further information, please email: reports@nie.edu.sg

Centre for Research in Pedagogy and Practice
National Institute of Education
1 Nanyang Walk
Singapore 637616
http://www.crpp.nie.edu.sg