Science and Mathematics Education Centre

Investigating the Effectiveness of a Professional Development Initiative for Lower Secondary Teachers in Indonesia

Titien Soewastiningsih Soebari

This thesis is presented for the Degree of

Doctor of Philosophy

of

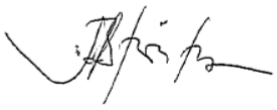
Curtin University

September 2012

DECLARATION

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

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

Signature: 

Date: 31 August, 2012

ABSTRACT

The overarching aim of this study was to evaluate the effectiveness of a teacher professional development initiative, designed to help lower secondary teachers to improve their teaching practice. To evaluate the effectiveness of the programme, a model of evaluation, influenced by three existing models, was developed. This holistic model was designed to suit the Indonesian context and included five distinctive phases. The first phase involved the collection of baseline data related to the teachers' classroom practices and students' perceptions of their learning experiences and attitudes. The second phase involved evaluating the teachers' views of the relevance and utility of the professional development programme. The third phase involved examining the extent to which the teachers' translated the knowledge and skills, imparted during the programme, into practice. The fourth phase involved examining whether there were changes in students' perceptions of their learning experiences and attitudes over the duration of the one year programme. Finally, the fifth phase sought to examine whether there were contextual factors that might promote or impede the translation of the professional development programme into practice. These five phases were used to guide the collection and analysis of data.

The sample for the evaluation involved different grain sizes. Quantitative data were collected from 2,417 students (drawn from 66 classrooms in 32 Indonesian lower secondary schools) who were present for both the pre-test and the post-test. Qualitative data, in the form of reflective journals, was collected from the 138 teachers who were involved in the professional development programme. In-depth interviews were collected from 33 teachers (from whose classes the quantitative sample was drawn). Finally, six of these 33 teachers formed the basis of case studies that involved the collection of observation and in-depth interview data.

Data collected from teacher reflective journals and interviews with teachers were used to evaluate the teachers' views of the relevance and utility of the teacher

professional development programme (Phase 2 of the evaluation model). The findings indicated that the teachers generally perceived the professional development experience positively, however, they were critical about two aspects of the delivery, these being the calibre of two of the four trainers and the tight time schedule required to complete the programme.

The findings suggested that there were differences in how the teachers translated the professional development ideas into their classroom practices (Phase 1 and 3 of the evaluation model). Two of the case study teachers, both of whom were teaching in urban schools, attempted to change their teaching practices to be more student-centred. Meanwhile, the teaching practices of the remaining four case study teachers (one from an urban school and three from rural schools) continued to be predominantly teacher-centred.

To examine the pre–post differences in students' perceptions of their learning experience and attitudes, MANOVA with repeated-measures were performed using the sample of 2,417 students. The results indicated that there were statistically significant (p<0.01) differences for the six of the seven WIHIC scales, but not the attitude scale, with all statistically significant differences showing an improvement for the post-test results. Despite the statistical significance for these WIHIC scales, the effect sizes were small (according to Cohen's criteria), suggesting that the success of MGMP professional development programme was limited.

The observations made at Phase 3 of the evaluation indicated that the MGMP Empowerment programme may have been more effective for teachers in urban schools than their counterparts in rural schools. A two-way MANOVA, with the WIHIC and attitude scales as the set of dependent variables and the two independent variables involving a two-level variable based on location (urban and rural) and two-level repeated-measure testing occasion (pre-test and post-test) indicated that, for six of the seven learning environment scales (the exception being the Involvement scale) and the attitude scale, a statistically significant interaction between testing occasion and location existed. For the WIHIC scales with a

statistically significant interaction, in all cases, a locality gap existed at the pre-test, which increased at the post-test. For these scales, students in urban classrooms scored higher for the pre-test and the post-test than their rural counterparts. A comparison of the effect size indicated that, the magnitude of the changes were greater for the students in urban schools than their rural counterparts.

The results indicated that there were three contextual factors that influenced the teachers' efforts to translate the ideas of the professional development programme into practice, these being, the instructional context, the lack of support provided by the school administrators and the nation-wide leaving examination. Whereas the case study teachers in urban schools felt that the status of the schools and the socio-economic status of the families of the students supported their willingness to try out the new ideas, the teachers from rural schools claimed that issues related to student discipline, large classroom sizes and inadequate resources and facilities hindered their attempts to implement the new ideas in their classrooms. All of the case study teachers indicated that there was inadequate practical support from either the principals or superintendents. Finally, the nation-wide leaving examination was perceived as the main hindrance for all six case study teachers. Although student-centred approaches have been mandated in the national curriculum, it appears that the high-stakes national examination forced the teachers to continue to use teacher-centred approaches, as these approaches were viewed as more effective.

The findings of this study have resulted in the development of a professional development model that brings together the field of learning environments with the evaluation of professional development suitable for use in Indonesia and it is anticipated that the findings will contribute to a better understanding of how teacher professional development in Indonesia can be improved. As such, the model can be used in the evaluation of future professional development programmes in Indonesia.

DEDICATED TO

Brian Mandegani Suprayogo and Colin Madani Suparoyogo

My boys

who have complained least during my busiest time.

Teguh Suprayogo

My husband, my best friend

for your love, patience, sacrifice and support.

ACKNOWLEDGEMENTS

My doctoral studies would not have been possible without the help and support of many individuals, whom I would like to acknowledge and extend to them my thanks.

First, my sincere gratitude goes to Associate Professor Jill Aldridge, my academic supervisor, for her patient assistance, support, guidance and encouragement throughout my doctoral programme. I am greatly indebted to her for her constructive suggestions and generous help.

I would like to thank Professor Barry Fraser, Director of Science and Mathematics Education Centre (SMEC), Curtin University, for serving as my associate supervisor. Without his support, I could not have come to this point.

My warmest thanks to my generous friends, Rosemary Patroni and Natasha Lee Kepert, for their valuable help with proofreading. Special appreciation also goes to all of the staff at SMEC, including Professor David Treagust, Dr. Tony Rickards, Kate Ala'i, Trudy Tanner, and others who have supported me and provided a warm working environment.

My deepest thanks to my dear husband Teguh Suprayogo and my wonderful sons, Brian Mandegani Suprayogo and Colin Madani Suprayogo for their love, patience, understanding and support, and for the sacrifices that they have made throughout my study in Australia. My deepest appreciation and recognition also go to my parents (may they rest in peace) and my relatives, whose love, prayers and support have helped me to overcome the difficulties that I have experienced during my studies.

Finally, I would like to convey my appreciation to the Indonesian government, and the Department of Education and Culture, for providing me with the opportunity and financial support to pursue my study in Australia.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

INTRODUCTION AND OVERVIEW

In the history of education, the professional development of educators has been given prominence in proposals and plans for educational reform and school improvement (Guskey, 2000). Historically, however, the evaluation of teacher professional development has been limited to exit surveys which, notably, fail to examine the impact of the professional development efforts. Despite the importance of professional development to reform efforts, there are serious concerns about the evaluation of the effectiveness of professional development practice with much literature related to descriptions of past failures (Corcoran, 1995; Guskey, 2002; Guskey & Huberman, 1995). The overarching aim of this study was to evaluate the effectiveness of a professional development programme, developed to improve lower secondary school teachers' practices in Indonesia. In this evaluation, these previous shortcomings are overcome by, firstly, modifying existing models used to evaluate professional development and, secondly, using a range of data, including student perception feedback, to help to evaluate the effectiveness of the professional development programme.

This chapter introduces this study using the following headings:

- The Indonesian Context (Section 1.1)
- Background of the Study (Section 1.2)
- Research Aims (Section 1.3)
- Significance (Section 1.4)
- Overview of the Thesis (Section 1.5)

1.1 The Indonesian Context

This section provides a brief overview of the geographical location of Indonesia (Section 1.1.1); an explanation of the education system in Indonesia (Section 1.1.2); and a description of the education reforms that have and are taking place in Indonesia (Section 1.1.3).

1.1.1 Geographical Location of Indonesia

The Republic of Indonesia, together with Malaysia, the Philippines and Thailand, is a country in South East Asia. The Republic of Indonesia (hence forth referred to as Indonesia) is made up of 13,466 islands with five main islands stretching across some 3200 miles of equatorial ocean. Indonesia is the fourth most populous country in the world after China, India and the US; with a population almost 237.6 million people. Approximately 58% of Indonesians live on Java, the smallest of the five main islands, which is considered to be the most populous island in the world. Despite family planning programmes, that have been implemented since 1960s, the population of Indonesia is estimated to grow to around 254 million by 2020 (Statistics Indonesia, 2010).

According to Purwadi and Mulyoatmojo (2000), Indonesia consists of distinct ethnic, linguistic and religious groups across its many islands. There are more than 300 ethnic groups who speak more than 742 different languages and dialects. The largest ethnic group is the Javanese, who comprise 42% of the population, and who are the politically and culturally dominant group. Despite its ethnic and linguistic diversity, Indonesia has a shared identity that is defined by a national language, Bahasa Indonesia, the official language of the Indonesian people. Bahasa Indonesia is the instructional language used in teaching and learning in all formal education and is the language of business, politics, national media and academia. Consequently, Bahasa Indonesia is spoken by nearly every Indonesian (Purwadi & Mulyoatmodjo, 2000).

The Indonesian government stipulates the rights of religious freedom through its constitution, however, there are only six religions that are recognised officially, these being, Islam, Protestant, Roman Catholic, Hindu, Buddhism and Confucianism. Despite not being an Islamic state, Indonesia is, according to the 2000 census (Statistics Indonesia, 2010), the world's largest Muslim-majority nation, with 86.1% of all Indonesians being Muslim.

Administratively, Indonesia consists of 33 provinces, with each province having its own political legislature and governor. The provinces are divided into regencies (*kabupaten*) and municipalities (*kotamadya*), which are further divided into districts (*kecamatan*), and village groupings (either *desa* or *kelurahan*). Furthermore, a village is divided into citizen-groups (Rukun-Warga) which are further divided into neighbourhood-groups (Rukun-Tetangga). Following the implementation of regional autonomy measures, in 2001, the regencies have become the key administrative unit and is responsible for providing most government services. The village level of administration is the most influential on a citizen's daily life, with all matters being handled through an elected *lurah* or *kepala desa* (village chief) (Statistics Indonesia, 2010).

1.1.2 The Education System in Indonesia

According to the Act of the Republic of Indonesia Number 20, 2003 (Depdiknas, 2003), education in Indonesia should involve a conscious and well-planned effort to create an appropriate learning environment and learning process. The Act claims that the purpose of education is to develop learners to their full potential by ensuring that they have opportunity to acquire spiritual and religious strengths, self-control, personality, intelligence, morals, noble character and skills that are required for his or her own sake, for the community, for the nation and for the state. The National Education system is based on Pancasila (the official philosophical foundation that forms the basis of the 1945 Constitution of the Republic of Indonesia). As such, the education system is rooted in the religious

values and national cultures of Indonesia (Depdiknas, 2003) and is purported to be responsive to the needs of an ever-changing world.

The Act states that the national education system functions to develop the capability, character and civilization of the nation by enhancing intellectual capacity and developing learner's potentials so that they are imbued with human values leading them to be faithful, pious and possess noble character. In addition, the education system serves to develop the potential of learners to ensure wealth, knowledge, competence, creativity and independence.

The education in Indonesia is controlled by two ministries; the Ministry of National Education of Indonesia (MONE- *Kementerian Pendidikan Nasional Republik Indonesia*) and the Ministry of Religious Affairs of Indonesia (MORA- *Kementerian Agama Republik Indonesia).* Public secular schools and non-Muslim private schools are the responsibility of the Ministry of National Education (MONE), while Islamic schools are administered by the Ministry of Religious Affairs (MORA).

Formal education in Indonesia consists of three levels, these being: (1) basic education (grade 1 to 6); (2) secondary education (grade 7 to 12); and (3) higher education. The education system is known as the "6-3-3-4" in which there are six years of primary school, three years of lower secondary school, three years of senior secondary school and four years of tertiary education.

In 2000, Indonesia began to decentralise the management of all government sectors. As a result of this decentralisation, the education system of Indonesia adopted a school-based management scheme (Dediknas, 2003). This scheme has seen the central government give autonomy in terms of management to schools in all areas, including the evaluation of teachers' works, which is now the responsibility of principals and superintendents.

At the World Education Forum in April, 2000, the government of Indonesia, along with other governments and organisations, made political commitments to achieve

basic education for all children. To this end, the Indonesian government has been implementing a Compulsory Basic Education programme (consisting of nine years) with the aim that all children aged between seven and 15 years will attend school. This compulsory education includes six years of primary and three years of lower secondary education. Under this policy, known as the Wajib Belajar Pendidikan Dasar (Wajar Dikdas for short), all citizens between the age of seven and fifteen years should have equal and ample opportunities to receive formal schooling (Depdiknas, 2003). To date, Indonesia has achieved compulsory basic education for all primary school level and, presently, approximately 94.4% of children within this age group (seven to 15 years) are enrolled in national primary schools (Yulaelawati, 2004).

1.1.3 Education Reforms in Indonesia

Since independence from the Dutch in 1945, the Indonesian government has implemented a range of efforts in a bid to provide quality education for its people. In the 1970s, education reform policies were largely centred on expanding human resources and, as a result, have achieved compulsory education at the primary school level (six years). In the 1990s, the educational reform initiatives were focused on improving the quality of compulsory education. The rationale for focusing on compulsory education was that, in spite of large population and extensive natural resources, Indonesia's education system was not highly developed. At the national level, these reform efforts included the expansion of compulsory education, enhancing science and technology, curriculum decentralisation, improving the quality of textbooks and teachers' guides, developing the effectiveness of in-service teacher training and promoting a conducive school and classroom environment (; Ministry of National Education, 2000; Southeast Asian Ministry of Education Organization, 1999). Despite these substantial efforts, however, the quality of education (as reflected in the students' outcomes in national examination) was not considered to be satisfactory (Umaedi, 2000).

According to Umaedi (2000), the reform efforts were unsatisfactory largely because they focused on improving the quality of 'inputs' in the education process. These inputs included the provision of text books, teaching and learning media, curriculum reform and new school buildings. Despite improving the quality of the inputs, however, a range of problems continued to persist, including poorly trained teachers and inadequate textbooks and materials. The results of these reform efforts suggest that improving educational inputs alone was insufficient to improve students' outcomes. It would also appear that the top-down policy approach, in which the educational reform efforts were macro-oriented, centralised and headquartered at the national education office in Jakarta, also contributed to the lack of success.

Based on these unsatisfactory student outcomes, the Indonesian government has recently changed the focus of its school improvement projects to ensure that the quality of learning improves students' outcomes. A critical issue, therefore, has been to raise the teacher qualification standards to ensure good quality teachers (Katharina, 2009). As such, current education reform is being driven by a vision that emphasises the principles of democracy, autonomy, decentralisation and public accountability. Reform efforts within the education sector are focused on enhancing its performance and providing a more equitable distribution of educational opportunities. These reform efforts have had a fundamental impact on the Indonesian national education system (Indonesian National Education System, 2003).

The present study evaluates the effectiveness of one of the teacher professional development programmes that was implemented as part of the most recent educational reform efforts in Indonesia.

1.2 Background of the Study

This section provides a background to the study that is organised into the following headings: Conceptual Framework (Section 1.2.1); Teacher Professional Development (Section 1.2.2) and the MGMP Empowerment as a Teacher Professional Development Programme (Section 1.2.3).

1.2.1 Conceptual Framework

According to Cohen, Manion and Morrison (2007), educational research is essentially concerned with exploring and understanding social phenomena that are educational in nature. As such, social phenomena pertains to formalised and/or spontaneously occurring social, cultural and psychological processes which could be termed as education. Because theoretical questions in education emerge from different conceptions and interpretations of social reality, different paradigms have evolved to determine the criteria according to which one would select and define problems for inquiry. Thomas Kuhn (1970) characterises a paradigm as an integrated cluster of substantive concepts, variables and problems together with corresponding methodological approaches and tools.

In the past, researchers in the social and behavioural sciences have engaged in what is known as paradigm wars (Johnson & Onwuegbuzie, 2004; Newman, Ridenour, Newman & DeMarco, 2003). In these wars, or debates, researchers of a positivist worldview (historically these are generally quantitative researchers) tend to criticise constructivists for being too subjective and unreliable (Guba & Lincoln, 1989, 1994; Tashakkori & Teddlie, 1998, 2003). Similarly those researchers with an interpretativist worldview (historically these are generally qualitative researchers) tend to criticise empiricists for being too reductionist. These debates are essentially between purists (Johnson & Onwuegbuzie, 2004) and tend to focus on the differences rather than similarities between the two positions (Onwuegbuzie & Leech, 2005). In addition, these purists have argued that paradigms and research methodologies cannot be separated or mixed (Howe, 1988).

Although, historically, there have been arguments against mixing research paradigms and research methodologies (Howe, 1988), more recently researchers have argued that it is time to adopt a third paradigm, that of pragmatism (Morgan, 2007; Onwuegbuzie & Leech, 2005). Pragmatists argue for the importance of integrating methods when it is appropriate (Onwuegbuzie & Leech, 2005). Mixed-method research, according to Creswell and Plano Clark (2010), has philosophical assumptions as well as methods of inquiry that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches during the research process. As a method, it focuses on collecting, analysing and mixing both quantitative and qualitative data in a single study or series of studies.

The present study involved a mixed-methods approach by combining multiple research methods to help to evaluate the effectiveness of a teacher professional development programme. The mixed-methods approach used in the present study involved collecting, analysing and integrating both quantitative and qualitative data in a sustained programme of inquiry (Creswell & Plano Clark, 2010). The methods of data collection and data analysis are detailed in Chapter 3.

1.2.2 Teacher Professional Development

According to Newmann, Rutter and Smith (1989), teacher professional development (often referred to as in-service training, teacher training or staff development) has historically, been thought of as any formally planned activity intended to advance individual and collective staff members' knowledge, skills or expectations in a bid to improve student learning. Teacher professional development is also defined as a process of continual intellectual, experiential and attitudinal growth (Richards & Farells, 2005). Guskey (2000) suggested that continuous teacher professional development is essential as the knowledge base in education is growing rapidly.

Professional development for teachers has been described in different ways. Friedman and Philips (2004), for example, indicated that teacher professional development has been often perceived as formal training courses linked to work or

gaining a qualification. However, an emerging paradigm has moved teacher professional development away from the practice of attending courses or training to the concept of lifelong or continuing learning (Middlewood, Parker & Beere, 2005). Middlewood et al. (2005) argue that teacher professional development is an ongoing process of reflection that articulates, with development planning to meet corporate, departmental and individual needs.

It is suggested that teacher professional development can be taken to represent processes that, whether intuitive or deliberate, individual or social, results in specific changes in the professional knowledge, skills, attitudes, beliefs or behaviours of teachers. As such, teacher professional development programmes are generally designed to alter the professional practice, beliefs or understanding of school staff towards an 'articulated end' (Griffin, 1983, p.2; Guskey, 2002). In most cases, the articulated end is improved student learning. It is anticipated that, improving principals' and teachers' knowledge and skills in management, leadership and teaching and learning will, in turn, affect students' achievement.

Given that high quality professional development will produce better-quality teaching in classrooms, which, in turn translates into higher levels of student achievement, the quality of teacher professional development is a critical issue that must be addressed (Richardson, 1996; Supovitz, 2001, Supovitz & Turner, 2000). Teacher professional development programmes are rarely evaluated in systemic ways (Goodall, Day, Lindsay, Muijs & Harris, 2005; Guskey, 2000). According to Guskey (2000), the evaluation of teacher professional development programmes has commonly involved recording activities or outlining the activities undertaken without consideration of their impact on learning or practice. Evaluation of professional development has generally involved feedback questionnaires designed to gauge the participants' enjoyment of the activity rather than the impact of the activity on broader outcomes (Broad & Evans, 2006). The question, then, is what model of evaluation that can be used to evaluate the effectiveness of a teacher professional development programme in terms of teacher learning and student learning. It is with this in mind that the present study was initiated to evaluate the

effectiveness of a teacher professional development in terms of teacher learning and students' learning experience.

1.2.3 The MGMP Empowerment Programme

In Indonesia, English is a compulsory subject from the first grade of lower high school and is officially taught throughout the secondary school. The teaching and learning of English is considered to be a priority for the Ministry of National Education as well as Indonesian students (Depdiknas, 2003). For Indonesian students, learning English serves two purposes; first, students are required to read English texts in their college years; and second, competence in the English language is regularly used as a determining factor in securing favourable positions in the job market (Lie, 2007).

According to Indrajati Sidi (2008), the English language is considered, by Indonesian students, to be one of the most difficult subjects. In Balitbangdiknas' (2009) report of student results on the nationwide leaving examination (the Ujian Nasional) the average score for English language was low and, in spite of the six years of English instruction, the outcomes were unacceptable. According to Balitbangdiknas (2009), these low scores suggest that the current teaching practices used in English classes need to be improved.

As a result, considerable financial resources were brought to bear to change the English curriculum (carried out in 2004 and 2006). It was strongly suggested that, in addition to changes in the curriculum, professional development for teachers was also required to bring about improvement in student outcomes (Shawer, 2010). These changes have included methods for assessing the new student learning standards and, most recently, on assisting teachers to better help their students to achieve these new standards (Mendiknas, 2006). To address this and, attesting to the government's strong commitment to improving the quality of English language teaching, professional development for teachers has been provided through various

projects, one of which is the Musyawarah Guru Mata Pelajaran (MGMP) Empowerment programme for English, that was evaluated in the present study.

The Musyawarah Guru Mata Pelajaran (MGMP) Empowerment programme (hereafter referred to as the MGMP Empowerment programme) focused on developing teaching and learning models that would enable teachers to prepare creative, critical and skilful students, with the goal of improving student achievement. Upon completion of this programme, teachers were expected to employ more student-centred approaches in their practices by applying various instructional innovations, such as Contextual Teaching and Learning (CTL), Student Active Learning and Problem-Based Learning, among others. The implementation of the MGMP Empowerment programme involved teachers' participation in workshops, action research and lesson study (peer planning and peer observation) to help them to change their teaching practices. Also included in the programme was the teaching of skills to integrate technology into their English classes.

The MGMP Empowerment program is referred to as an 'in-on' system. That is, the programme was scheduled for 16 weeks over one academic year, during which there were two one-week in-service cycles, and two seven-week on-service cycles. Each MGMP induction cycle began with a week-long residential in-service component that was held at a provincial centre. This in-service cycle involved master teachers (guru pemandu), all of whom were selected from the regencies (kabupaten or kota), based on their merit, by the Education Quality Assurance Office. It was anticipated that the master teachers would attend the in-service workshops at the provincial level and, when they return to their regencies, they would spread the induction across an increasing number of teachers. Each regency sent between three and six master teachers to the in-service component, depending on the number of schools in the district.

During the in-service component, the master teachers worked on a number of key elements, one of which was a content analysis of the topics that they taught, involving the identification of major concepts and skills and planning of teaching

strategies. Also included in the in-service component were micro teaching sessions, during which instructors provided demonstration lessons to the participating master teachers that they were expected to deliver, in a similar fashion, to their colleagues.

After the in-service component, the master teachers returned to their respective schools for the on-service component of MGMP Empowerment cycle. During this on-service training, the master teachers each invited 17 English teachers who taught in schools located near to their own school to attend the weekly meetings. Therefore, each of these on-service groups involving 17 teachers and three master teachers in one on-service training location were known as MGMP clusters. This on-service training component consisted of meetings held once a week over 14 weeks, spread across two semesters. These weekly meetings were used to discuss how the teachers implemented the ideas suggested during the teacher professional programme and for master teachers to provide teaching demonstrations and micro teaching activities to the teachers. These meetings also provided opportunities for teachers to examine and discuss issues related to the dissemination of new educational policies, curriculum implementation, the development of lesson plans and worksheets and the analysis of test items. A diagrammatical representation of the organisational structure of the MGMP empowerment programme and an example of the programme used are provided in Appendix A and B, respectively.

The second residential in-service component took place mid-semester. During this time, the master teachers evaluated their on-service experiences and prepared for the second half of the semester (in a similar fashion to the training held at the beginning of the semester). Following the second in-service component, the master teachers returned to their schools for the remainder of the semester, thus completing one full cycle of MGMP programme.

Financial resources were required to implement the MGMP activities. The costs involved included payment for presenters, hiring of facilities, transport cost for the

participants, and training materials, including stationeries. The funding for these various activities was provided by the Ministry of National Education.

The project goals of the MGMP Empowerment programme were to develop professional activities aimed at improving teaching practices in ways that met the standards of educational services (within the framework of the Indonesian national education quality assurance). More specifically, the project goals (USAID, 2003) were to:

- encourage teachers to improve their ability and skill to plan, implement, and evaluate teaching and learning activities in their school;
- discuss problems faced by teachers to implement their daily responsibilities and to propose solutions in accordance with the characteristics of the subject matter, teachers, school conditions and communities;
- provide teachers with opportunities to share information and experiences about the implementation of the curriculum;
- provide teachers with opportunities to express their ideas through MGMP meetings to improve their profession;
- develop some cooperation with other institutions to develop conducive, effective, and enjoyable teaching and learning process.

The present study, due to the time restrictions, focused on evaluating the effectiveness of only the in-service component of the MGMP Empowerment programme. As such, this study only included the master teachers. In particular, this study involved examining the impact of the MGMP Empowerment programme on how these master teachers (henceforth referred to as teachers) changed their teaching practices in ways that were suggested by the programme. Throughout of this thesis, form this point forwards, the terms teacher professional development and the MGMP Empowerment programme are used interchangeably to refer to in-service programme.

1.3 Research Aims

The overarching aim of the present study was to examine the impact of a one-year teacher professional development programme (the MGMP Empowerment programme). To do this, several research questions were delineated. As a first step, to ensure that evaluation was meaningful to the Indonesian context, the first research objective was:

To develop a model suitable for evaluating teacher professional development in the Indonesian context.

Because part of the evaluation involved the administration of two surveys, it was necessary to select, modify and translate suitable instruments. To ensure the reliability and validity of the surveys, the second research objective was:

To modify, translate and validate Indonesian versions of two surveys (one to assess students' perceptions of the learning environment and one to assess students' attitudes towards their English classes) for use at the lower secondary school level in Indonesia.

To determine whether differences exist between students' perceptions and attitudes before and after the MGMP Empowerment programme, the third research objective was:

To examine the effectiveness of the MGMP Empowerment programme in terms of:

a) changes in teaching practice
b) students' perceptions of their learning experiences
c) students' attitudes towards English classes.

To examine whether the MGMP Empowerment programme was differentially effective for teachers in urban and rural schools, the fourth research objective was:

To investigate whether the MGMP Empowerment programme was differentially effective for teachers in urban and rural schools in terms of:

a) students' perceptions of their learning experiences
b) students' attitudes towards English classes.

To evaluate whether there are contributing factors influencing the implementation of the MGMP Empowerment programme, the fifth research objective was:

To examine whether contextual factors exist that might promote or impede changes in teaching practice.

1.4 Significance

It is anticipated that this study will have theoretical, methodological and practical implications. As such, it is hoped that the results of this study will benefit researchers, government policy makers, teacher professional development providers and teachers.

Theoretically, the present study has developed a professional development evaluation model that brings together the field of learning environments with the evaluation of professional development suitable for use in Indonesia. It is anticipated that the results of this study will provide a model that can be used in the evaluation of future professional development programmes in Indonesia.

The present study provides methodological implications to the field of learning environments through its contribution and translation of widely-used questionnaires for use in English classes in Indonesia. The Indonesian versions make available practical, economical and reliable questionnaires that can be used to help to determine whether, the students' experiences have changed as part of the evaluation of a professional development programme.

Practically, the results of the present study have the potential to provide insights to a variety of stakeholders, including teachers and the professional development providers who were central in this evaluation. The results could provide useful feedback that can be used to improve to design future professional development programmes in Indonesia.

1.5 Overview of the Thesis

This chapter introduced the thesis, providing a context for the study and an overview of the aims of the research. The chapter also provided a theoretical framework for the study and outlined the significance of the study.

Chapter 2 provides a review of literature relevant to the present study. The review includes the research behind current practices in teacher professional development, as well as current thinking about the importance and role of evaluating teacher professional development programmes. The chapter reviews pertinent literature related to the field of learning environments, highlighting past research and important findings. Also reviewed in this chapter is literature related to student attitudes.

Chapter 3 describes the model to evaluate the teacher professional development programme used for the present study and how this was modified from existing models. This chapter goes on to provide a rationale for using mixed-methods approach and describes the procedural aspects of the present study, including the nature and size of the sample, the selection, modification and translation of surveys, data-collection and approaches to analyse of the quantitative and qualitative data.

Chapter 4 describes the statistical analyses and results of the validation of the two surveys used as part of the evaluation of the professional development programme (one to assess students' perceptions of the learning environment and the other to assess students' attitudes). This component of the study ensured that the surveys

were valid and reliable in the Indonesian context, to provide confidence in the results.

Chapter 5 provides a detailed account of the data analysis and results which relate to research objectives three to five. This chapter involves evaluating the MGMP Empowerment programme using the model developed for use in Indonesia. The chapter reports both the quantitative and qualitative results which examine the effectiveness of the professional development in terms of changes in teachers' teaching practices, students' perceptions of their classroom experiences and students' attitudes towards English classes. Finally, this chapter examines the contextual factors that serve to promote or impede the translation of the professional development programme into practice.

Chapter 6 provides a discussion and conclusions related to the results of this study. The chapter also includes discussion of the limitations and potential implications of the findings of the study. The significant theoretical, methodological and practical contributions of the research are highlighted. The thesis concludes with suggestions for possible future research.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

Research in the field of teacher professional development has suggested that effective professional development has positive influences on teachers' knowledge, beliefs, skills and dispositions which, in turn, influences their instruction in ways that are productive for student learning (Darling-Hammond & Sykes, 2003; Garet, Porter, Desimone, Birman & Yoon, 2001). These studies have inferred that understanding the links between teacher professional development and student outcomes, in making judgements about the effectiveness of professional development, is important. It was noted, however, that relatively little systematic research has been conducted to examine the effects of professional development on improvements in teaching practices and students' outcomes. My study sought to overcome this problem by investigating the effectiveness of a professional development initiative for lower secondary English teachers in terms of changes in teachers' practice, students' perceptions of their learning experience and their attitudes towards the subject. This chapter reviews literature pertinent to my study and provides a background for achieving the aims of my study (see Chapter 1). The review is organised into the following subheadings:

- Teacher Professional Development: Theories and Models (Section 2.2),
- Effective Teacher Professional Development (Section 2.3),
- Evaluating Teacher Professional Development (Section 2.4),
- Learning Environments (Section 2.5),
- Student Attitudes (section 2.6), and
- Summary (Section 2.7).

2.2 Teacher Professional Development: Theories and Models

Fullan (2001) defined teacher professional development as the sum total of formal and informal learning pursued and experienced by the teacher in a compelling learning environment under conditions of complexity and dynamic change. Teacher professional development was defined by Hassel (1999), as the process of improving staff skills and competencies that are needed to produce outstanding educational results for students. Since teacher professional development is described in many different ways, it is interesting to note that its definitions are often categorised by paradigm. According to Broad and Evans (2006), teacher professional development is generally categorised into four paradigms, these being, the deficit paradigm, professional growth paradigm, educational change paradigm and problem solving paradigm.

The deficit paradigm refers to professional development that targets a lack of skills or knowledge and often views teachers as empty vessels that are to be filled (Gall & Renchler, 1985). Based on this deficit paradigm, professional development consists of one-shot workshops which aim to provide teachers with skills and knowledge. In this paradigm, teacher change has been directly linked with planned professional development activities (Clarke & Hollingsworth, 2002).

The professional growth paradigm characterises development as a more self-directed activity arising out of the teacher's interests and needs (Feiman-Nemser, 2001). In this paradigm, teacher professional development involves individual professional learning and development within a community of support. This paradigm highlights the importance of an individual teacher's needs and strengths.

The educational change paradigm views professional development as having a focus on bringing about change (Fullan, Hill & Crevola, 2006; Warren-Little, 2001). This paradigm refers to professional development efforts that are designed to facilitate change whereby the major goals of professional development programmes are to

change teachers' classroom practices, their attitudes and beliefs and the learning outcomes of students (Guskey, 2002).

The problem solving paradigm links professional development to making improvements to address issues that have been identified by teachers such as student achievement needs (Ball & Cohen, 1999; Joyce & Showers, 2002; McLaughlin & Zarrow, 2001). In this paradigm, teacher professional development opportunities should increase teachers' awareness of new learning practices and their knowledge to solve issues that are happening in their own classroom (Joyce & Shower, 2002).

The following sections discuss literature pertaining to different models of teacher professional development that are related to more traditional perspectives (Section 2.2.1) and the current thinking about how effective professional development should be designed from a more contemporary perspective (Section 2.2.2).

2.2.1 Models of Teacher Professional Development

Traditionally, educators have had a somewhat narrow view of professional development, perceiving it as special events that occur at discrete times throughout a school year (Monahan, 1996). As a traditional approach, teacher professional development was based on a training or deficit paradigm implying a shortfall in teachers' skills and knowledge (Guskey, 1986). This perception of professional development has contributed to the predominate set of beliefs, held by teachers, that professional development is distinct and separate from the daily work that they do in the school.

Under these more traditional perspectives, professional development programmes have been treated as 'one shot' workshops, aimed at providing teacher mastery of prescribed skills and knowledge. This model of professional development has involved 'sit and get' courses or programmes. In some countries, the same approach to professional development is viewed as graduate courses, in which

teachers enrol to attain an advanced degree or to move ahead to a higher salary scale. This view is also supported by some government policies that required teachers to accumulate a certain number of professional development hours or credits each year in order to retain their jobs and professional certification (Guskey, 2000).

In most traditional models of teacher professional development, courses are designed to improve professional practices and are developed by external authorities. Such courses are often delivered on a 'one off' basis and held outside of the setting in which they will be implemented. This model of professional development is known as an outside-in professional development model which draws, for the most part, upon knowledge that has been generated by outsiders for teachers to use in their practice (Hoban, 1997, 2002). This model supports the traditional role of educational researchers who regularly work outside of schools and produce knowledge for teachers inside schools as represented in Figure 2.1.

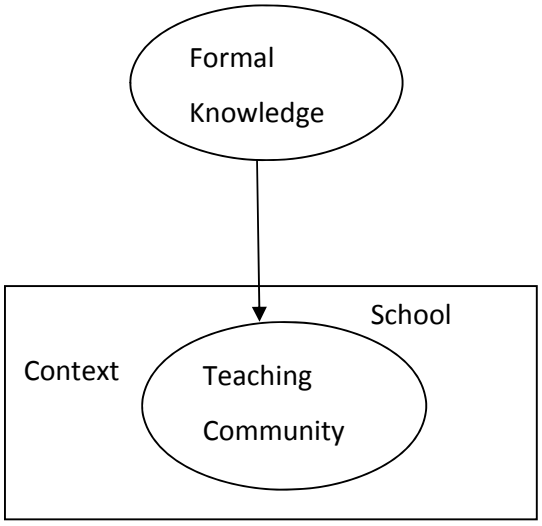


Figure 2.1. Representation of the Outside-In Professional Development Model

Gaible and Burns (2005) identified the outside-in model as the standardised teacher professional development model. This model represents a centralised approach,

involving workshops and training sessions. As noted by Hatton (2000), the rationale for the outside-in or standardised models is that teachers, as a community, tend to reproduce existing practice and need to be informed by educational theory to provide alternative perspectives on teaching and learning. As the role of educational researchers is to generate new knowledge about teaching and learning, it is common to use formal knowledge as the content for professional development programmes to inform teachers about new ideas.

The strengths of the outside-in models for teacher learning are two-fold in that it: 1) provides new ideas about teaching practice that are often beyond teachers' existing experiences; and 2) is a convenient and economical way to disseminate new knowledge because teachers can gather at a venue and be provided with content in a brief period of time (Hoban, 1997, 2002). In addition, the topic of a workshop may match with teachers' existing beliefs or focus on an area that the teachers are interested in learning more about. Moreover, it would appear that a major benefit of training workshops is the opportunity for teachers to interact with colleagues both within a school and from other schools. Sparks and Loucks-Horsley (1990, p. 241) noted that such a model was beneficial because it promoted practice that is deemed to be worth replicating and claimed that "…there are behaviours and techniques worthy of replication by teachers in the classroom... teachers can change their behaviours and learn to replicate behaviours in their classroom that were not previously in their repertoire".

Similarly, Guskey (1998), together with Gaible and Burns (2005), note that all workshops and presentations that were offered outside of regular school hours or on specially designated professional development days, could be highly appropriate. Such professional development is considered to be particularly effective when educators need to obtain information about new programmes, new instructional approaches, or changes in school policies and regulations. It is important to recognise, however, that all workshops and presentations must be accompanied by appropriate follow-up activities.

Despite the strengths of the outside-in model, this model has been criticised. The literature suggests there are many limitations of the outside-in model and there is a need for a richer repertoire of professional development practices and models of delivery for teachers (Brown-Easton, 2004; Feiman-Nemser, 2001; Guskey, 2000; Lieberman, 1995; Lieberman & Miller, 2000, Lieberman & Wilkins, 2006). Researchers, including Fullan (2003), highlight that the outside-in models are often ineffective because: 1) the teachers' existing practices are not taken into consideration; 2) the influential conditions of the school context are not considered; and, 3) in many cases there is insufficient follow-up to support and promote change.

Lieberman (1995), amongst others, described the outside-in model as 'direct teaching' with the teacher as the passive recipient of new knowledge. Lieberman (1995, p. 591) denounces the outside-in model stating that "teacher professional development as a transferable package of knowledge to be distributed in bite-sized pieces needs radical rethinking; it implies a limited conception of teacher that is out of step with the current research and practices".

In addition, it is argued that there are five components identified as key considerations of effective professional development that are not taken into account in the outside-in model (Joyce & Shower, 2002). These five key components (that maximise the probability of achieving the desired effects) are the:

- *exploration of theory* including discussions or readings to assist participants to understand the ideas being presented;
- *demonstration or modelling of the ideas* in the form of videotape or live in a setting that can be integrated with explanations of theory;
- *practice of the required skill under stimulated conditions* to approximate the workplace that needs to occur twenty or twenty five times depending on the complexity of the tasks;
- feedback on the required skill from peers or experts following practice;

- *coaching in the workplace following initial training* involving a supportive community of teachers to provide assistance and collegiality during the learning process.

Other researchers, including Fullan and Stiegelbauer (2007), Johnson (2001), and Lovitt and Clarke (1998), have provided convincing evidence of the failure of professional development activities based on this model. Their findings imply that this model is ineffective because teachers were generally passive participants during and after the completion of professional development programmes. Furthermore, this model of professional development is not congruent with notions of engaged, self-directed professionals and situated, social and constructed workplace learning (Webster-Wright, 2009).

2.2.2 Current Views of Professional Development

Past reviews related to the ineffectiveness of traditional models of teacher professional development has led to a surge of research and suggestions on how to improve the planning and delivery of teacher professional development programmes (Elmore, 2004; Fullan, 2001; Hirsh, 2006; Loucks-Horsley, Katherine & Peter, 2003; Sparks, 2002). As a result, there has been a shift in focus from traditional professional development models, in which the emphasis is on changing the teachers (who remain as passive participants in the process), to more complex processes that involve teachers playing an active role in shaping their professional growth through reflective participation (Clarke & Hollingsworth, 2002).

Based on these new perspectives there is a general consensus that the role of teachers in their own professional development should be central (Fullan & Hargreaves, 2002). As such, there is trend which recognises that, as professionals learn, their knowledge bases become broader and more specialised and that their capacity to increase and extend their knowledge and skills also grows (Brown-Easton, 2004; Joyce & Showers, 2002). Hoban (1997, 2002) coined the term inside-in professional development to describe a model of professional development that

draws upon teachers' existing knowledge that they have generated from their school experiences and encourages them to reflect on and explore their own ideas, based on these understandings. These models are believed to be consistent with workplace programmes, which encourage teachers to take responsibility for their own professional development. The inside-in model is represented below in Figure 2.2.

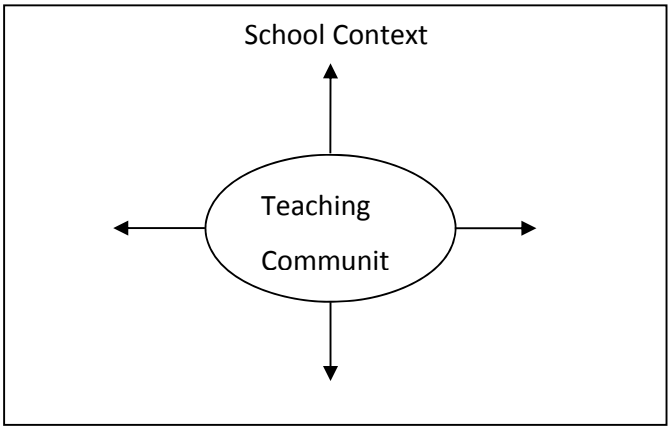


Figure 2.2 Representation of the Inside-In Professional Development Model

A key aspect underpinning the inside-in model is personal reflection, which is used to assist teachers to become aware of how they understand their practice (Baird, 1992). Sparks and Loucks-Horsley (1990), Louck-Horsley, Katherine & Peter, 2003) described four teacher professional development models as being inside-in models, these being: the Individually Guided Staff Development Model; the Observation/ Assessment Staff Development Model; the Development/Improvement Staff Development model; and the Inquiry Staff Development Model. Each of these models is described below.

The Individually Guided Staff Development Model encourages individual teachers to take responsibility for their professional learning. An assumption of this model is that individuals are motivated by being able to select their own learning goals and the means by which they accomplish those goals. A belief that underlies this model is that self-directed development empowers teachers to address their own problems and that, doing so, creates a sense of professionalism. Examples include

teachers experimenting with new instructional strategies that they have devised, visiting other classrooms, devising personal growth plans and studying new instructional materials.

The Observation/Assessment Staff Development Model involves teachers working in pairs to provide an alternative view on classroom instruction. An assumption of this model is that instructional practices are improved if a colleague or other person observes a teacher's classroom and provides feedback. Having someone else in the classroom to view instruction and to provide feedback or reflection also is a powerful way to impact classroom behaviour. Observers also learn as they view their colleagues in action. Examples of programmes include peer coaching, clinical supervision and teacher evaluation.

The Development Improvement Staff Development Model typically involves assessing current practices and determining a problem, the solution of which will improve student outcomes. This model involves teachers working as a team to develop curriculum or programme design, development or review, to implement new instructional strategies or to solve problems. Examples of this model include a collaborative model which involves teachers working as a team to develop their own school-based curriculum or school improvement programme.

Inquiry Staff Development Model involves teachers working individually or collectively, in learning by selecting a problem and determining an action to take. Teachers formulate questions about their own practice and pursue answers to those questions. Inquiry involves the identification of a problem, data collection (from the research literature and classroom data), data analysis, and changes in practice followed by the collection of additional data. The inquiry can be done individually or in small groups. This model is built on the belief that the mark of a professional teacher is the ability to take reflective action. The popular example of this model is action research which involves a continuing spiral consisting of four major parts: planning, acting, observing, and reflecting.

According to Hoban (1997, 2002) the strengths of inside-in teacher professional development models are that they: 1) encourage teachers to be responsible for their own learning; 2) encourage teachers to value their ideas as they are based on evidence collected within their context; and 3) are consistent with a constructivist perspective of learning that encourages teachers to be reflective practitioners.

In contrast to the outside-in model, Rochsantiningsih (2006) identified that in-side in model of teacher professional development (i.e. action research as professional development) provides a powerful means of professional growth as it was closely related to the teachers' teaching situations. As such, this model creates changes and innovations in teachers' practices that lead to more effective classroom activities and teacher-student interaction.

Gaible and Burns (2005) described the inside-in model as a self-directed teacher professional development model. As such, teachers are involved in initiating and designing their own professional development and in sharing materials, ideas, challenges and solutions with colleagues. Gaible and Burns (2005) consider that, although teachers should be encouraged to participate in ongoing, self-motivated learning and self-directed activities, they should not be used as the primary means of providing teacher professional development as they are likely to interpret their experiences based on the way they already frame their practice. This is more likely to occur if teachers work in isolation and do not collaborate with other teachers to provide them with alternative perspectives.

The two models, outside-in and inside-in, both have strengths and limitations. Past literature has considered ways to involve aspects of both outside-in and inside-in professional development models to become inside/outside models. These professional development models draw upon both the knowledge that teachers have generated from their experiences and the knowledge of others to promote a community of discourse (Cochran-Smith & Lytle, 1993). Lieberman and Miller (2000) for example, indicated that direct teaching of new ideas through courses, workshops and conferences has merit, particularly to develop awareness of new

research or methodologies, and job-embedded professional development practices through peer-coaching, mentoring, action research, planning teams and critical friendships offer strong opportunities to implement practice.

Inside/outside models involve two aspects of teacher learning: personal reflection by the participants; and the introduction of alternative views to provide participants with different perspectives on teaching and learning. Figure 2.3 below presents a diagram that depicts the inside/ outside model.

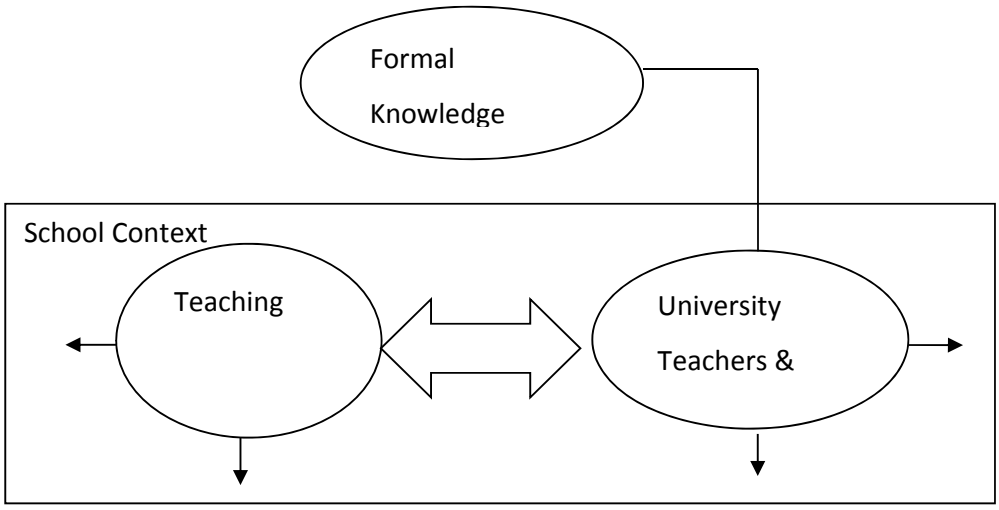


Figure 2.3 Representation of Inside/Outside Model

There are two major strengths of the inside/outside models. First, the models involve an interaction between the knowledge that teachers have generated from their experiences and the alternative perspectives which may be introduced by university- based teacher educators, parents or students. Second, professional development programmes based on this model are usually long term involving regular meetings of the participants. Possible limitations, however, centre on the needs of the different interest groups because of their various agendas and, the amount of time and effort that it takes to establish and maintain such community discussions (Cochran-Smith & Lytle, 1993; Guskey, 2000; Hoban, 1997, 2002).

2.3 Effective Teacher Professional Development

With the increasing emphasis on education that is linked to technology and workforce change and increasing responsibility to promote students' critical thinking and problem solving, there are new demands on teachers' professional competence to facilitate students' learning (Fullan, 2003). As a result of these changes there is also a need to shift teachers' professional duties from transferring discipline knowledge to teaching for understanding (Loucks-Horsley et al., 2003; Sparks, 2002). Clearly, to prepare teachers to teach to more challenging standards, effective professional development is needed. Richards and Lockhart (2000, p. 40) pointed out the importance of the best professional development practices:

> We need to find the best ways for helping teachers to explore their practice … that practice involves exploring the relationship between the individual teacher's thinking and acting within the four walls of the classroom and the relationship between what the teacher does in the classroom and how this reproduces and/or transforms values and social ideals in society.

Research has indicated that, regardless of teacher's diverse motivations to participate in teacher professional development programmes (e.g. for certification or contractual agreements), most teachers engage in professional development activities because they want to become better teachers (Fullan, 2003; Huberman, 1995). For the vast majority of teachers. Guskey (2000) purports that teachers participate in teacher professional development programmes to become better teachers which will, in turn, enhance student learning outcomes. It is my experience, however, that in some countries, teachers attend professional development to fulfil contractual obligations and/or government mandates to ensure that they retain their employment. It is interesting to note that in previous studies, most teachers defined their professional success in terms of their students' behaviour and activities, rather than in terms criteria related to teacher professional improvement (Fullan, 2003; Fullan & Hargreaves, 2002). Fullan and Miles' (1992) also reported that the attraction, to teachers, to take part in teacher professional

development was their belief that professional development would expand their knowledge and skills, contribute to their growth and enhance their effectiveness with students.

Further to these statements, Griffin (1993) defines professional development programmes as systematic efforts to bring about change in the classroom practices of teachers, in their attitudes and beliefs, and in student learning outcomes. Guskey (2000) proposed that professional development involves processes and activities designed to enhance the professional knowledge, skills, and attitudes of educators so that they might, in turn, improve the learning of students. Mathison (1992) defined effective teacher professional development (the term 'in-service teacher education' was used) as an experience that would:

- guarantee institutional support;
- increase teachers' knowledge and skills;
- change classroom practices;
- change students' learning experiences; and
- promote changes in an organisation.

Increasingly, studies have examined the relationships between professional development and improvements in student learning, in ways that are more systematic (see for example Guskey & Huberman, 1995; Lieberman, 1995). Guskey and Sparks (2002, p.73) propose a model of the relationship between professional development activities and improvements in student learning claiming that:

> The premise of the model is that the quality of professional development is influenced by a multitude of factors. Those believed to have the most immediate and direct influence, however, can be classified in three major categories: content characteristics, process variables, and context characteristics.

In this model, the content characteristics refer to the 'what' of professional development (new knowledge, skills and understandings that are the foundation of any professional development effort). Process variables relate to the 'how' of professional development (type, forms and planning of activities), and context characteristics refer to the 'who', 'when', 'where' and 'why' of professional development (organisation, system and culture in which professional development takes place).

Spark and Hirsch (2000) propose three transformative ideas that are important to the design of effective teacher professional development, these being: results-driven education; systems thinking; and constructivism. Professional development involving results-driven ideas, are judged not only by a participant's attendance at a professional development activity, but rather by the extent to which changes in teaching practices, considered to be valuable to students, are observed. Professional development, involving system thinking ideas, can be judged on the movement from a view of discrete events to the interconnectedness of elements (e.g. collaboration and collegiality). And finally, professional development design based on a constructivist basis integrates the general concepts of learners and learning.

In Thompson's (2003) review of evidence-based research, related to the impact of professional development on teachers' practice and student learning, several points of convergence were found. According to Thompson (2003), these studies highlight the features of teacher professional development that are most likely to contribute to improved classroom practice and student performance including:

- a focus on subject matter learning;
- a link to curricular materials and assessments;
- the promotion of 'coherence' and 'active learning'; and
- extended activities that permit more active learning, and promote collective participation.

Congruent with these findings, Fishman, Marx, Best and Tal (2003) developed a model of teacher professional development that linked teacher's learning with changes in the knowledge, beliefs and attitudes of teachers, leading to the acquisition of new skills, concepts and processes that are related to the work of teaching. This professional development model (depicted in Figure 2.4) takes into consideration that teacher learning within a system of professional development involves an interactive relationship between the student change or learning that is represented by various forms of assessment, and through practical experiences of classroom enactment.

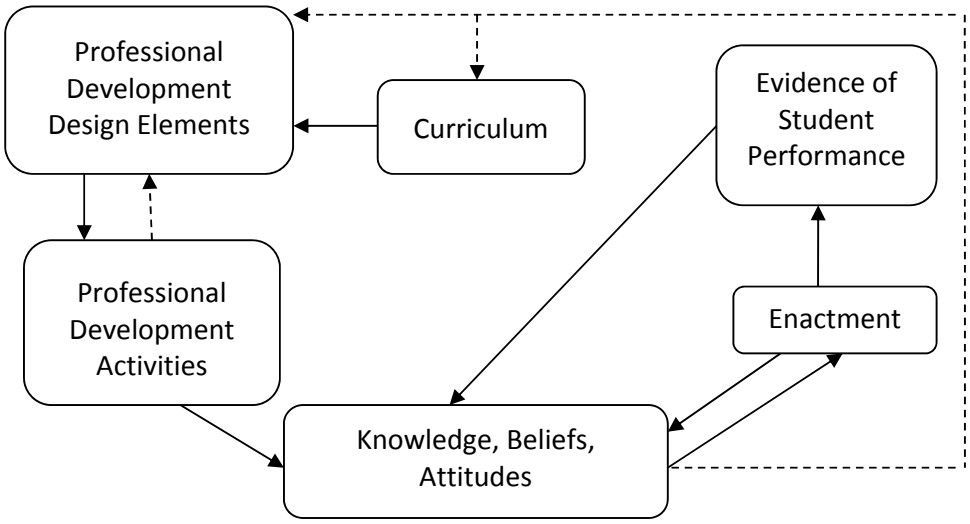


Figure 2.4 Fishman et al.'s (2003) Model of Teacher Professional Development

The model, represented in Figure 2.4, shows teachers' knowledge, beliefs, and attitudes to be the aspects of teacher cognition that are affected by participation in professional development programme. Following Richardson's (1996) stance, that teachers' knowledge, beliefs and attitudes are formed interactively with classroom enactment, each aspect is likely to influence the other. This model of teacher professional development, therefore, links enactment directly, and reciprocally, to teachers' knowledge, beliefs and attitudes. Student performance also influences teachers' knowledge, beliefs and attitudes, and are mediated, intuitively, as teachers look to their students for feedback about the instruction. Using this model,

then, teacher professional development elements can be modified to increase their impact on teachers' knowledge, beliefs, attitudes and subsequent enactments.

In more recent literature, there has been an increasing emphasis on effective teacher professional development and the learning environment (Darling-Hammond, 1998; Loucks-Horsley, et al., 2003; Loucks-Horsley & Matsumoto, 1999; Love, 2002). Stigler and Hiebert (2004) suggest that, in designing a teacher professional development, a focus on teaching must avoid the temptation to consider only the superficial aspects of teaching (e.g. the organisation, tools, curriculum content, and text books). The cultural activity of teaching, the ways in which the teacher and students interact and the delivery of the subject, can be more powerful than the curriculum materials alone. With a similar emphasis on the importance of the learning environment, Weiss and Pasley (2004) indicate that, for the continuing professional growth of teacher, professional development activities should reflect elements of high-quality instruction with clear, explicit learning goals and a supportive but challenging learning environment.

The MGMP Empowerment programme, evaluated in my study, drew upon and extended the inside/outside model of teacher professional development (presented in Figure 2.3). The MGMP Empowerment programme, referred to as an 'in-on system', involved two components, these being an in-service cycle and an on-service cycle. In the in-service cycle, the teachers were involved in Training of Trainer's workshops led by instructors. During the workshops, the teachers were presented with formal knowledge that has been generated in other contexts (such as learning theory). During the on-service cycle, the teachers spread the induction (formal knowledge) across an increasing number of teachers during on-service workshops. At the same time, the teachers shared their knowledge and skills with their colleagues by involving them in action research and lesson studies at their school sites. This involved weekly meetings of English teachers to work together and share insights to generate teaching strategies to foster student learning. (For further details related to the MGMP Empowerment Programme, refer to Section 1.2.1.)

2.4 Evaluation of Teacher Professional Development

Despite the recognition of the importance of evaluating teacher professional development and the need for formative information to develop effective programmes, to date most professional development programmes are not evaluated in a systematic way (Broad & Evans, 2006). The evaluation of professional development has consisted, on the whole, of documenting the event (Killion, 2002). Such documentation has focused, typically, on the participants' reactions to whether they liked the training, the quality of the presenter, whether their own comfort needs were met, and how useful they thought that the information might be to them (Goodall, et al., 2005; Guskey, 2000). Sparks & Hirsch (2002) described this form of evaluation as a happiness quotient. While most teacher professional development has aimed to change teaching practice and to enhance student outcomes, until recently, attempts at evaluation of such programmes have rarely included direct measures of the impact of the professional development on either of these aspects.

It is notable that, over the past ten years, there has been an emerging interest in evaluating professional development. Guskey (2002, p. 8) describes the four reasons for the growing interest:

- A better understanding of the "dynamic nature" of professional development.
- Recognition of professional development as "an intentional process".
- The need for better information to guide reform efforts.
- Increased pressure for "accountability".

Despite awareness of the need for, and interest in, evaluating professional development, little progress has been made (Mizel, 2003). According to Mizel (2003), the field of staff development needs better evaluation both to improve the

effectiveness of teachers' learning experiences and to produce credible evidence that will bring together more support for professional development.

Current literature recommends that the evaluation of teacher professional development practices should be based on five general standards. These are outlined by Stiggins (2001) as:

- Standard 1: Evaluation must be derived from clearly specified targets and outcomes.
- Standard 2: The results of evaluation must be used to develop further learning or plan action.
- Standard 3: Appropriate evaluation methods that reflect the desired outcomes and evidence are required.
- Standard 4: Evaluation must provide a sufficiently varied amount of evidence to allow confidence in the evaluation and forward planning.
- Standard 5: Evaluation must be designed, developed and used in such a manner as to eliminate bias.

Guskey and Roy (1995) suggested a minimum set of guidelines to be established for evaluating professional development programmes. These guidelines might consider that the evaluation:

- be ongoing;
- procedures be explicit and public;
- be informed by multiple sources of data;
- use both quantitative and qualitative data;
- focus on all levels of the organization;
- be considerate of participant's time and energy;
- results be presented in forms that can be understood by all programme participants and patrons.

William and Burden (1994, 2000) has categorised the evaluation of professional development into summative, formative or illuminative. Summative evaluation is carried out to judge the overall effectiveness of the programme, and occurs at the end of the programme. Formative evaluations are used to modify or improve a programme, and are conducted during the course of the programme, usually at regular intervals, so as to cater to the needs of the participants. Formative evaluations are, therefore, ongoing in nature, and seek to inform, improve and direct the project, rather than simply assess its impact. Illuminative evaluation involves the using a variety of sources, acknowledging multiple perspectives and taking into account the background and culture of the context. In all case, the evaluator should be integral to the day to day working of the project, and ensure that data is used to assist in decision making and to guide implementation.

Kennedy (1998) explained that, in evaluating a teacher professional development programme, the concern should not only consider evaluating the outcome but should also consider the processes of the teacher professional development programme. That is, the evaluation should start at from the implementation stages; from identification of a problem to the eventual diffusion of the innovation. Similarly, Tribble (2000) noted that evaluation should attempt to interpret the impact of the professional development programme in a bid to understand the reasons for its success or failure.

A review of literature revealed that there are three predominant models that have been developed to guide the evaluation of teacher professional development programmes. The first, developed by Mathison (1992), (presented in Figure 2.5) identifies generic aspects of professional development for teachers. That is, the professional development process is the provision of an experience, the implementation of new knowledge and skills by teachers and followed by changes in student learning. Mathison (1992) outlined each aspect of the model of professional development by suggesting questions that should be asked at each stage and the types of data collection methods to be used. These are outlined briefly below.

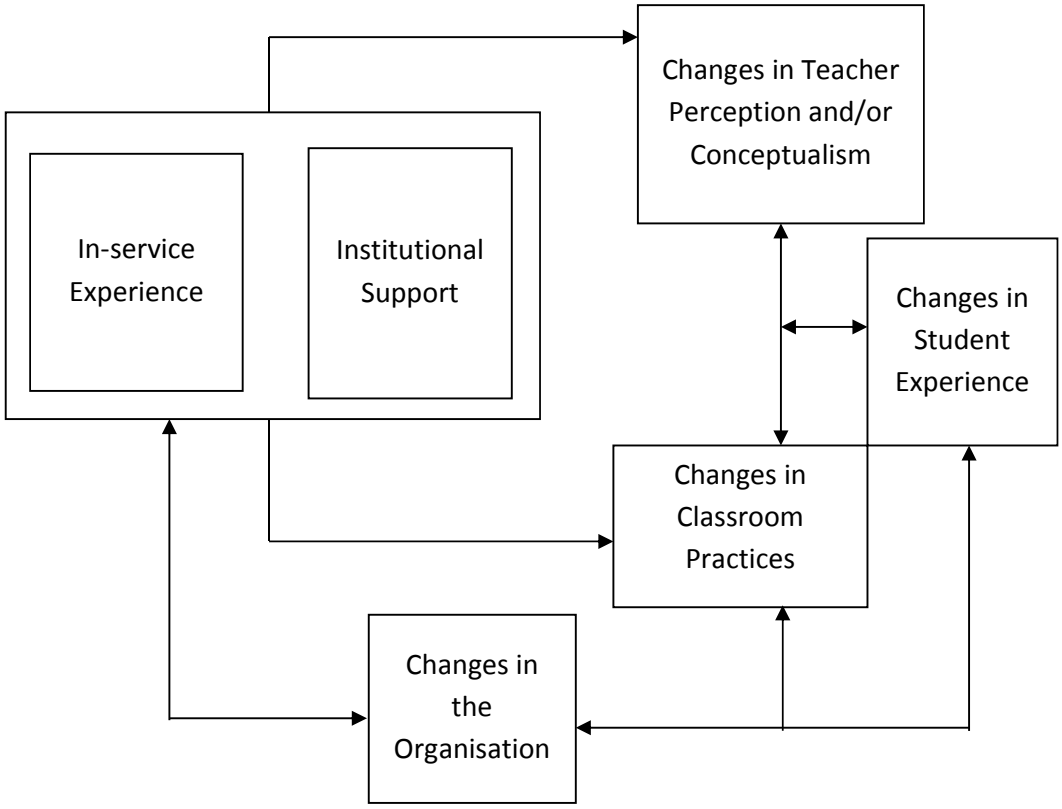


Figure 2.5 Mathison's Model of Professional Development Evaluation

- The in-service experience evaluates the professional development experience in terms of the relevance of the teacher professional development programme to the teachers. In this aspect, teachers' reactions and responses provide data about what is happening during the programme and what is crucial to the ongoing development of the programme.

- Institutional support assesses the extent to which school networks provide ongoing assistance to teachers in their efforts to change, as well as increasing collegiality. The evaluation in this aspect focuses on whether the networks within school exist and how successful are the networks in providing desirable support for teachers during the implementation.
- Change in teacher perceptions evaluates teachers' perceptions and conceptions before, during and after professional development programme.

It is anticipated that, by participating in the teacher professional development programme, the teachers increase their knowledge and hold more positive attitudes towards the subject that they teach.

- Changes in classroom practices focuses the evaluation on changes in classroom practices as the teachers adopt the knowledge from the professional development programme. The data from this aspect of evaluation provides feedback for improvement for future teacher professional development programme.

- Changes in student experience evaluates on the extent to which students' engagement in learning activities, enjoyment towards the lesson and achievement.

- Changes in organisation, assesses the degree to which changes in the organisation support the professional development for teachers. This aspect of evaluation also evaluates contributing factors that might influence the implementation.

Table 2.1 provides evaluation questions for each aspect of professional development model that was represented in Figure 2.5 and the strategies by which data might be collected.

Table 2.1 Mathison's (1992) Evaluation of Teacher Professional Development

Aspect of Professional Development (PD)	Evaluation Questions	Data Collection Strategies
In-service (PD) Experience	• Is the content correct? • Are teachers treated as professionals and is their sense of professionalism enhanced? • Is the experience relevant to the world of teaching? • Does the in-service provide good role models for teachers? • Does the in-service accommodate differences among teachers? • Are multiple purpose of in-service reasonably well served?	Questionnaires Test of Content Interviews Participant Observation
Institutional Support	• Is it provided? • What form does it take? • What institutional barriers exist? • Is it effective?	Interviews Observations
Change in Teacher Perceptions and/ Conceptualisations	• Do teachers think differently about X? • How do they think about X? • Do teachers feel good about themselves and their work?	Interviews Questionnaires
Change in Classroom Practices	• Do teachers do noticeably different things in their classrooms? • Are they able to successfully do different things in their classrooms?	Interviews Classroom Observations
Change in Student Experience	• Are students exposed to new/different ideas/ activities? • How do students feel about their experiences? • Are students doing different things? • Is student learning positively affected by changes?	Classroom observations Interviews Tests Questionnaires
Change in Organisation	• What effect does in-service have on organisational structures? • Effect of organisational climate? • Effect on organisational operations? • Does organisation contribute to/impede change seen as desirable by teachers?	Interviews Questionnaires Participant Observations

In another model, Fishman et al. (2003) recognised that a chain of evidence, from multiple points across the model of teacher professional development (presented in Section 2.3 and Figure 2.4) is important. This model of professional development

was used to develop a model for the evaluation of professional development programmes (illustrated below in Figure 2.6). The evaluation of professional development using this model involves six aspects, these being: standards, evidence of student performance; professional development design; evaluation of professional development; observation of classroom teaching; and evaluation of student performance. Fishman et al. (2003) outlined each of the aspects to be evaluated each of which are described below.

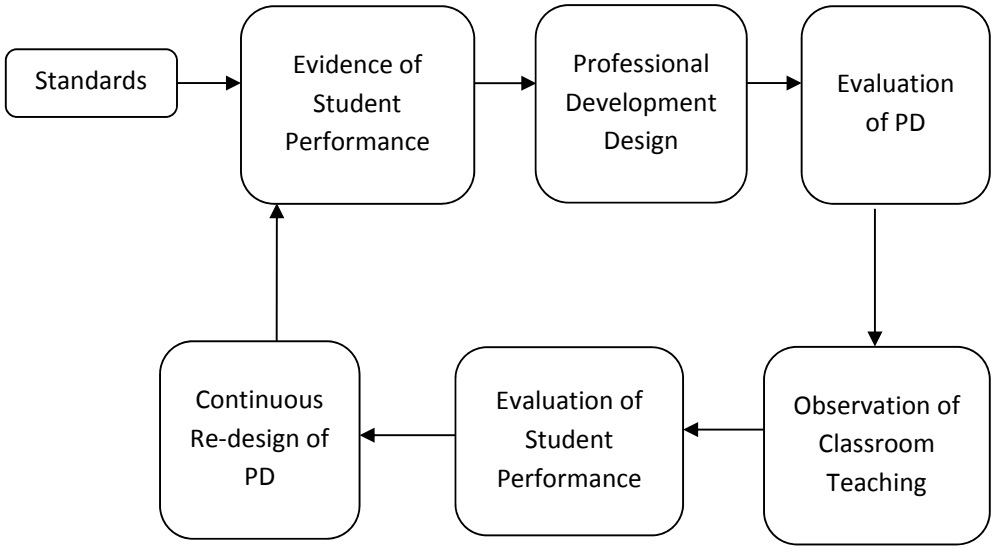


Figure 2.6 Fishman et al.'s (2003) Model of Professional Development Evaluation

- Standards refer to those standards from which the evaluation begins and involves education standard documents and curricula based on the materials of the standard documents.

- Evidence of Student Performance seeks evidence related to current student performance from student artefacts, classroom behaviours, or pre-test, related to the curriculum units.

- Professional Development Design involves a focus on evaluating the design of the professional development programme and it is intended to help teachers to acquire the knowledge necessary to successfully enact the

curriculum units. This aspect also focuses on areas where student performance may be improved.

- The Evaluation of Professional Development evaluates teachers' attitudes towards the professional development programme and their beliefs about whether they gain new knowledge and skills during the professional development programme. This component of the evaluation usually involves the use of interviews.

- Observation of Classroom Teaching examines changes in classroom practices by looking for evidence of teaching behaviours that match what is taught during the professional development, and what teachers say that they will try to do in post-professional development interviews. Interviews are also conducted after observations to ask the specific questions about what they have done in the classroom.

- The Evaluation of Student Performance evaluates changes in student learning with respect to the professional development activities.

According to Guskey and Sparks (2002), sound evaluation of professional development should provide information that is meaningful, and sufficiently reliable to use in making thoughtful and responsible decisions about the effects of the professional development processes. Based on this principle, Guskey (2000) developed a model for evaluating teacher professional development in which he identified five critical levels, these being: (1) Participants' Reaction; (2) Participants' Learning; (3) Organization Support and Change; (4) Participants' Use of New knowledge and Skills; and (5) Student Learning Outcomes.

- At Level 1, Participants' Reaction, the participants' initial satisfaction with the experience is assessed. This level aims to provide information that can help to improve the design and delivery of programmes or activities. Guskey (2000)

points out that a positive reaction from a participant is usually a necessary prerequisite to higher- level evaluation results.

- At Level 2, Participants' Learning, there is a focus on assessing the knowledge and skills that participants gained as a result of professional development. Analysis of the information gathered at this level provides a basis for improving the content, format and organisation of the programme or activities.

- At Level 3, Organisation Support and Change, the focus is on the organisational characteristics and attributes necessary for successful implementation of the professional development ideas. From a systemic point of view, information at this level helps to document the organisational conditions that accompany success or describe those that might explain the lack of significant improvement.

- At Level 4, Participants' Use of New Knowledge and Skills, focuses on the evaluation of whether participants use their new knowledge and skills on the job. The information at this level rests in the clear specification of indicators that reveal both the degree and quality of the implementation.

- At Level 5, Student Learning Outcomes assesses the impacts of professional development on students. In some cases, the information related to student learning outcomes is used to estimate the cost effectiveness of professional development.

Table 2.2 outlines the five levels of evaluation by Guskey (2000) along with the questions to be addressed in each level, the instruments to be used and the aspects to be measured in each level.

Table 2.2 Guskey's (2000) Professional Development Evaluation

Evaluation Level	Questions to be Addressed	Instruments	What is Measured
1. Participants' Reactions	Did they like it? Was their time well spent? Did the materials make sense? Will it be useful? Was the leader knowledgeable and helpful? Was the room the right temperature and were the refreshments tasty?	Questionnaire/ survey Focus group Interview Personal log/ reflective journal	Initial satisfaction with the experience
2. Participants' Learning	Did participants acquire the intended knowledge and skills?	Simulations Demonstrations Participants' reflections Participants' portfolios Case study analyses	New knowledge and skills of participants
3. Organisation Support and Change	Were sufficient resources made available? Were problems addressed quickly and efficiently? Was implementation advocated, facilitated, and supported? Were successes recognised and shared? What was the impact on organisation? Did it affect organisational climate and procedures?	Questionnaires Minutes from follow-up meeting Focus group Interview Participants' portfolios District and school records	The organisation's advocacy, support, accommodation, facilitation, and recognition
4. Participants' Use of New Knowledge and Skills	Did participants effectively apply the new knowledge and skills?	Questionnaires Interviews Participants' reflection Participants' portfolios Direct observation	Degree and quality of implementation
5. Student Learning Outcomes	What was the impact on students? Did it affect students' performance or achievement? Did it influence students' physical or emotional wellbeing? Are students more confident as learners? Is student attendance improving? Are dropouts decreasing?	Student/ school records Questionnaires Interviews with student, parents, teachers, and/ or administrators Participant's portfolios	Student learning outcomes: Cognitive Affective Psychomotor

A review of literature related to professional development of teachers, shows an increasing tendency to advocate the use of a range of models to evaluate the effectiveness of teacher professional development. A multi-site evaluation model, for example, was conducted by Abell, Lannin, Marra, Ehlert, Cole, Lee, Rogers and Wang (2007) to investigate the effectiveness teacher professional development opportunities for science and mathematics in the United States. This model extended Guskey's (2000) model of evaluation through the creation of individual projects while providing an evaluation picture of the overall programme. Abell et al. (2007) provided an illustrative evaluation by highlighting the importance of understanding the contexts of the professional development. Through two case comparisons, they demonstrated how a case-based context can be used to inform their interpretation of outcomes data.

In another evaluation, conducted by Hanley, Maringe and Ratcliffe (2008), a change transition model was used to explore the processes of a three-phase professional development programme to support teachers' expertise in science teaching. The four features of the change transition model (trigger, vision, conversion and maintenance) were used as a framework for the analysis of the data. The trigger layer was concerned with the need for change, in terms of opportunities, threats to the individuals and organisation, the crises the organisation faces and its needs for the future. The vision layer involved establishing the future development of the organisation by articulating a vision and communicating the vision effectively. The conversion layer involved an implementation phase, including persuading and converting people to commit to the vision. The maintenance and renewal layer involved the management of mid-term change. A clear identification of mid-term outcomes is needed as progress is made towards the final intended product.

A hybrid evaluation model of teacher professional development programme was used by Hahs-Vaughn, Zygouris-Coe and Fiedlers (2007) to evaluate the Florida Online Reading Professional Development (FOR-PD) project in the United States. The hybrid evaluation model incorporated: 1) the logic model (Harris, 2001); 2) the four levels of evaluating training programmes (Kirkpatrick, 1976); 3) Guskey's model

(Guskey, 2000); and 4) the five pillars of quality online education (Mayadas, Bourne & Moore, 2002). The combination of four models for evaluating online learning helped to ensure that critical perspectives were included in the evaluation.

The professional development evaluation models and studies that were presented in this section provided a range of recommendations aimed at coherence and increased sophistication in evaluation. Some focused upon outcomes (Fisman et al., 2003; Mathison, 1992), while others emphasised context variables (Abell et al., 2007), impacts of the delivery models (Hahs-Vaughn et al., 2007) and links between the professional development and student learning (Guskey, 2000).

My review of literature highlights the potential for combining elements of different models to tailor a model for professional development to suit a specific context. My study drew on and extended these past studies by developing an evaluation model to suit the needs of the Indonesian context. In doing so, my study paid explicit attention to indicators of evaluation and included students' perceptual measures to evaluate the effectiveness of the MGMP Empowerment programme.

2.5 Learning Environments

There is compelling evidence that the classroom learning environment strongly influences student outcomes, so much so that it should not be ignored by those wishing to improve the effectiveness of schools (Fraser, 2001). While the research base is rich with the effect of learning environment on student outcomes few, if any, past evaluations of professional development have involved a learning environment perspective as a component. My review of literature led me to include students' perceptions of the learning environment, as an important aspect of the evaluation of the effectiveness of the MGMP Empowerment programme.

The psychosocial learning environment refers to the tone, ambience or atmosphere created in a classroom. With respect to the learning environment, Moos (1979, p. vii) coined the terms "social climate" and "personality of the environment", to

describe the learning environment. This section reviews literature related to theories that have influenced the field of psychosocial learning environments and how this has led educational researchers to study the learning environment as an alterable educational variable which influences students' cognitive and affective outcomes. This section discusses the history of the field of learning environments (Section 2.5.1), instruments for assessing the learning environment (Section 2.5.2), past research within the field of learning environments (Section 2.5.3) and the use of learning environment as a source of process in the evaluation of educational innovations (Section 2.5.4).

2.5.1 History of the Field of Learning Environments

As discussed in chapter 1, the notion of a learning environment existed as early as 1936 when Lewin proposed that personal behaviour is a result of the interaction between the individual and his/her environment. To this end, he developed the formula $B=f(P,E)$ in which behaviour (B) is a result of the interaction between the person (P) and environmental factors (E). Building on Lewin's theory, Murray (1938) proposed his Needs-Press theory in which an individual's behaviour is affected internally by characteristics of personality (needs) and externally by the environment itself (press). Press, according to Murray (1938) has a directional tendency with properties not obtainable by the sum of the parts of the environment or situation.

Stern, Stein and Bloom (1956) further proposed that the same environment can be perceived differently by different entities, namely, individuals, groups and external observers of the environment. Independent of this thread of research, Stern, Stein and Bloom (1956) also pointed to measurements of educational environments as decisive components for prediction and successful learning manipulation. Hunt (1975), Stern (1970) and Fraser and Fisher (1983) proposed the notion of person-environment fit in which an individual is likely to achieve better outcomes (cognitive and affective) if the environment is more closely matched to the environment that they would prefer.

In 1981, Walberg proposed a nine-factor model of educational productivity in which student outcomes are co-determined by such variables as the quantity and quality of instruction, the psychosocial environments of the school/class, the home, the peer group and the mass media (Fraser, Walberg, Welch & Hattie, 1987; Walberg, 1981). In their research, carried out to examine whether correlations exist between student outcomes and the various factors proposed in the nine factor model, Fraser, Walberg, Welch and Hattie (1987) found that the psychosocial environment was a strong predictor of both achievement and attitudes even when a comprehensive set of other factors were held constant.

Moos (1991) proposed that the different characteristics of all human environments can be classified into the three broad dimensions of Relationship Dimension, Personal Development Dimension and the System Maintenance and System Change Dimension. The Relationship Dimension assesses "the extent to which people are involved in the setting, the extent to which they support and help each other, and the extent to which they express themselves freely and openly" (Moos, 1979, p. 14). The Personal Development Dimension assesses "the basic directions along which personal growth and self enhancement tend to occur in the particular environment" (Moos, 1976, p. 331). Finally, the System Maintenance and System Change Dimension assess the "extent to which the environment is orderly and clear in its expectations, maintains control and responds to change" (Moos, 1979, p. 16). These dimensions co-exist in all human environments and have been used extensively by researchers in the construction of learning environment instruments (Fraser, 1998, 2007; Walker, 2006) and the classification of individual scales.

In the 1960s, the first two psychosocial learning environment instruments were developed independently of each other: the Learning Environment Inventory (Walberg & Anderson, 1968); and the Classroom Environment Scale (Moos & Trickett, 1987). Since this time, much work has been done to conceptualise the learning environment and to assess students' perceptions of their educational environments (Fraser, 2007). The development of an International Journal dedicated to this field of study, *Learning Environments Research* (Fraser, 1998), as

well as books such as *Studies in Educational Learning Environments* (Goh & Khine, 2002) and *Outcomes-Focused Learning Environments* (Aldridge & Fraser, 2008), among others, have helped to inform the worldwide educational community of the importance of this area of research.

The following sections review literature related to the development of instruments to assess the learning environment (Section 2.5.2), the types of past research that have been conducted within the field of learning environments (Section 2.5.3) and the evaluation of educational innovations (Section 2.5.4).

2.5.2 Instruments for Assessing Learning Environments

Over the past 40 years numerous, valid and reliable instruments have been developed to enable teachers and researchers to assess students' perceptions of the learning environment. These questionnaires have been used at different educational levels and have been translated and used in different countries. This section provides a brief description of nine historically-significant and contemporary instruments, these being: Learning Environments Inventory (Section 2.5.2.1); Classroom Environment Scale (Section 2.5.2.2); Individualised Classroom Environment Questionnaire (Section 2.5.2.3); My Class Inventory (Section 2.5.2.4); College and University Classroom Environment Inventory (Section 2.5.2.5); Science Laboratory Environment Inventory (Section 2.5.2.6); Questionnaire on Teacher Interaction (Section 2.5.2.7); Constructivist Learning Environment Survey (Section 2.5.2.8); and What is Happening In this Class? Questionnaire (Section 2.5.2.8).

A summary of the nine instruments is provided in Table 2.3. The table shows the name and scales of each instrument, the level at which it was developed for use, the number of items in each scale, and the classification of each scale according to Moos (1974) scheme for classifying human social environments.

Table 2.3 Overview of Nine Learning Environment Instruments

Instrument	Level	No. of Items	Scale Classified According to Moos' Scheme		
			Relationship Dimension	Personal Development Dimension	System Maintenance and Change Dimension
Learning Environment Inventory (LEI)	Secondary	7	Cohesiveness Friction Favouritism Cliqueness Satisfaction Apathy	Speed Difficulty Competitiveness	Diversity Formality Material Environment Goal Direction Disorganisation
Classroom Environment Scale (CES)	Secondary	10	Involvement Affiliation Teacher Support	Task Orientation Competition	Organisation Rule Clarity Teacher Control Innovation
Individualised Classroom Environment Questionnaire (ICEQ)	Secondary	10	Personalisation Participation	Independence Investigation	Differentiation
My Class Inventory (MCI)	Primary	6-9	Cohesiveness Friction Satisfaction	Difficulty Competitiveness	
College and University Classroom Environment Inventory (CUCEI)	Higher Education	7	Personalisation Involvement Student Cohesiveness Satisfaction	Task Orientation	Innovation Individualisation
Science Laboratory Environment Inventory (SLEI)	Upper Secondary Higher Education	7	Student Cohesiveness	Open-Endedness Integration	Material Environment Rule Clarity
Questionnaire on Teacher Interaction (QTI)	Secondary	8-10	Helping/Friendly Understanding Dissatisfied Admonishing Leadership Student Responsibility and Freedom Uncertain Strict		
Constructivist Learning Environment Survey (CLES)	Secondary	7	Personal Relevance Uncertainty	Critical Voice Shared Control	Student Negotiation
What Is Happening In this Class (WIHIC)	Secondary	8	Student Cohesiveness Teacher Support Involvement	Investigation Task Orientation Cooperation	Equity

Adapted with permission from Fraser (1998a)

2.5.2.1 Learning Environment Inventory

The Learning Environment Inventory (LEI) was developed and validated as part of the evaluation and research related to Harvard Project Physics (Walberg & Anderson, 1968). The final version of the inventory consisted of 105 statements, descriptive of school classes typical of that era. The 15 scales included Cohesiveness, Friction, Favouritism, Cliqueness, Satisfaction, Apathy, Speed, Difficulty, Competitiveness, Diversity, Formality, Material Environment, Goal Description, Disorganisation and Democracy. The items are presented in a cyclic order and the response scale involved the four alternatives of Strongly Disagree, Disagree, Agree or Strongly Agree. Example statements include 'All students know each other very well' (Cohesiveness) and 'The pace of the class is rushed' (Speed). Whilst some of the scales included in this questionnaire might still be applicable in today's settings, many are suited to a more traditional, teacher-centred learning environment.

2.5.2.2 Classroom Environment Scale

The Classroom Environment Scale (CES) was developed by Rudolf Moos at Stanford University (Moos & Trickett, 1987), as a result of extensive research that involved perceptual measures of a variety of human environments, including, psychiatric hospitals, prisons, university residences and work milieus (Moos, 1974). The final published version of the CES includes nine scales, namely, Involvement, Affiliation, Teacher Support, Task Orientation, Competition, Order and Organisation, Rule Clarity, Teacher Control and Innovation. There are 10 items in each scale with a True-False response format. Two example items are 'The teacher takes a personal interest in the students' (Teacher Support) and 'There is a clear set of rules for students to follow' (Rule Clarity). Although some scales continue to be useful, many are not pertinent for classrooms in more student-centred classrooms.

2.5.2.3 Individualised Classroom Environment Questionnaire

The Individualised Classroom Environment Questionnaire (ICEQ) assesses dimensions which distinguish individualised classrooms from traditional ones. The published version of the ICEQ (Fraser, 1990; Rentoul & Fraser, 1979) has 50 items

with 10 items in each of five scales, namely, Personalisation, Participation, Independence, Investigation and Differentiation. Each item is responded to using a five-point frequency scale of Almost Never, Seldom, Sometimes, Often and Very Often. Example items include 'The teacher considers students' feelings' (Personalisation) and 'Different students use different books, equipment and materials' (Differentiation). The ICEQ has been utilised in a number of studies and been found to be valid and reliable; including a study to examine the relationship between student classroom environment perceptions and their attitudes in Brunei (Asghar & Fraser, 1995). This important instrument was one of the first to consider the learning environment created in the more student-centred classroom. Although my study sought to examine, using students' perceptions, the extent to which the professional development programme changed teachers' practices to more student centred approach, the reliability of the ICEQ was not considered to be as robust (when translated and used in non-western contexts) as the instrument that I selected (described later in Section 2.5.2.9).

2.5.2.4 My Class Inventory

The My Class Inventory (MCI) is a simplified version of the LEI, and was developed for use with children aged 8–12 (Fraser, Anderson & Walberg, 1982; Fisher & Fraser, 1981; Fraser & O'Brien, 1985). Although initially developed for use with primary school students, the MCI has been found to also be useful for research involving secondary school students who experience reading difficulty. To make the instrument more manageable for younger children, the items were simplified to make them easier to read, the number of items and scales were reduced and the response format was reduced to a Yes–No format. Goh, Young and Fraser (1995) subsequently used a three-point response format of Seldom, Sometimes and Most of the Time to provide a more meaningful response format. The final version has 35 items within five scales, namely, Cohesiveness, Friction, Satisfaction, Difficulty and Competitiveness. Example items include: 'Children are always fighting with each other' (Friction) and 'Children seem to like the class' (Satisfaction).

The MCI has been used successfully in a number of countries, including, Brunei Darussalam (Majeed, Fraser & Aldridge, 2002), the United States (Scott Houston, Fraser & Ledbetter, 2008; Sink & Spencer, 2005) and Singapore (Goh, Young & Fraser, 1995). In Brunei Darussalam, the MCI was used among 1565 lower secondary mathematics students, and was found to have satisfactory factorial validity (Majeed, Fraser & Aldridge, 2002). In the US, two independent studies, one involving a sample of 2835 students in grades 4 to 6 (Sink & Spencer, 2005) and one involving 588 grade 3 to 5 students (Scott Houston, Fraser & Ledbetter, 2008), both found the MCI to have a satisfactory psychometric properties. Although these studies suggest that the MCI was suitable for use with younger students, this instrument assessed dimensions such as competitiveness, friction and difficulty (Fraser, 1989), which were not centrally relevant to my study.

2.5.2.5 College and University Classroom Environment Inventory

To fill the void of research in the area of classroom environment research at the higher education level, the College and University Classroom Environment Inventory (CUCEI) was developed (Fraser & Treagust, 1986; Fraser, Treagust & Dennis, 1986). The CUCEI was developed for use in tutorial classes (up to 30 students) and includes the seven scales of Personalisation, Involvement, Student Cohesiveness, Satisfaction, Task Orientation, Innovation and Individualisation. The response format involves a four-point rating scale of Strongly Agree, Agree, Disagree and Strongly Disagree. Example items include 'Activities in this class are clearly and carefully planned' (Task Orientation) and 'Teaching approaches allow students to proceed at their own pace' (Individualisation). While the CUCEI was created for the small classroom environment typical at a university, it has found applicability at the secondary level as well (Logan, Crump, & Rennie, 2006; Nair & Fisher, 2001).

2.5.2.6 Science Laboratory Environment Inventory

The Science Laboratory Environment Inventory (SLEI) was developed by Fraser, Giddings and McRobbie (1995) to assess the unique setting of science laboratory classes. The SLEI has seven items in each of five scales, namely, Student Cohesiveness, Open-Endedness, Integration, Rule Clarity and Material Environment

(Fraser et al., 1995; Fraser & McRobbie, 1995; Fraser, McRobbie & Giddings, 1993). The response format involves a five-point frequency scale consisting of Almost Never, Seldom, Sometimes, Often and Very Often. Example items are 'I use the theory from my regular science class sessions during laboratory activities' (Integration) and 'We know the results that we are supposed to get before we commence a laboratory activity' (Open-Endedness).

The SLEI has been used in studies around the world. When the SLEI was originally field tested, the sample involved 5447 students from six different countries, these being, the USA, Canada, England, Israel, Australia and Nigeria (Fraser, 1998). In the United States, Lightburn and Fraser (2007) used the SLEI with a sample of 761 high school biology students to examine the effectiveness of using anthropometry activities. The study reported the validity of the SLEI as well as the positive influence of anthropometry activities on the learning environment. In Korea, the SLEI was translated and used to examine whether perception differences in the learning environment existed for three science streams (Fraser & Lee, 2009). The study involved 439 high school science students and reported the factorial validity of the Korean version of the SLEI. The study also reported that students in different science streams perceived the learning environment differently.

2.5.2.7 Questionnaire on Teacher Interaction

Past research has shown that the interactions of the teacher can influence student outcomes. The Questionnaire on Teacher Interaction (QTI) was developed in the Netherlands by Wubbels, Creton and Hooymayers (1992) to evaluate students' and teachers' perception of interpersonal teacher behaviour. The QTI was based on Leary's (1957) work of Interpersonal Diagnosis of Personality. The theoretical model maps interpersonal behaviour using an *influence* dimension (Dominance – Submission) and a *proximity* dimension (Cooperation – Opposition) (Wubbels & Brekelmans, 2005; Wubbels & Levy, 1993). These dimensions are represented in a coordinate system divided into eight equal sectors which are Leadership, Helping/Friendly, Understanding, Student Responsibility/Freedom, Uncertain,

Dissatisfied, Admonishing, and Strict behaviour (Wubbels et al., 1992). The response format involves a five-point rating scale that ranges from Never to Always.

The QTI has been translated into different languages including Standard Malay, Korean and Indonesian and cross-validated at different grade levels in the USA (Wubbels & Levy, 1993), Australia (Fisher, Henderson & Fraser, 1995; Henderson, Fisher & Fraser, 2000), Singapore (Goh & Fraser, 1998; Quek, Wong & Fraser, 2005), Korea (Kim, Fisher & Fraser, 2010; Lee, Fraser & Fisher, 2003), Brunei (Scott & Fisher, 2004) and Indonesia (Fraser, Aldridge & Soerjaningsih, 2010). Although the QTI has been used in Indonesia in a previous study (Fraser, Aldridge & Soerjaningsih, 2010) and shown to be valid and reliable in a range of secondary school contexts, its' focus on teacher-student interactions made it unsuitable to evaluate the professional development programme, which involved a range of other factors related to student-centred learning.

2.5.2.8 Constructivist Learning Environment Survey

The Constructivist Learning Environment Survey (CLES) was developed to assess the extent to which a particular classroom's environment is consistent with constructivist epistemology, to help teachers to reflect on their epistemological assumptions and to reshape their teaching practice (Taylor, Dawson & Fraser, 1995; Taylor, Fraser & Fisher, 1997). To the constructivist, meaningful learning is an active process of constructing rather than acquiring knowledge, and instruction is a 'process of supporting that construction rather than communicating knowledge' (Duffy & Cunningham 1996, p. 171). The CLES has six items in each of five scales, these being, Personal Relevance, Uncertainty, Critical Voice, Shared Control and Student Negotiation. The response format involves a five-point frequency scale of Almost Never, Seldom, Sometimes, Often and Almost Always. The CLES was the first learning environment instrument to order the items in scales rather than cyclically to provide students with contextual cues, thereby improving the reliability of the instrument (Taylor, Fraser & Fisher, 1997). Two example items are 'I learn that Science has changed over time' (Uncertainty) and 'It's okay for me to express my opinions' (Critical Voice).

The reliability and usefulness of the CLES has been reported in numerous studies. The CLES was used in a cross-national study involving Taiwan and Australia (Aldridge, Fraser, Taylor & Chen, 2000), in which it was administered to 1081 students in 50 classes in Australia and a Mandarin version of the CLES was administered to 1879 students in 50 classes in Taiwan. The study reported sound psychometric properties of the CLES when used in both countries. In addition, the results indicated that Australian students perceived a more constructivist learning environment than their Taiwanese counterparts.

The CLES has been used in several studies in the United States. Peiro and Fraser (2009) modified and translated the CLES into Spanish and administered Spanish and English versions to 739 K-3 science students. Analysis of the data supported the validity of this modified version when used with young children. Johnson and McClure (2004) validated a 20-item version of the CLES. Nix, Fraser and Ledbetter (2005) used this 20-item version of the CLES with 1079 students in 59 science classes and reported sound psychometric properties for the revised version. In South Africa, Aldridge, Fraser and Sebela (2004) used the CLES with 1864 grade 4 to 6 mathematics students. The results reported the sound factorial validity and internal consistency of the CLES when used with this group of students.

Despite the value of the CLES, particularly in evaluating educational innovations in a range of countries, the focus of the CLES is on constructivist notions such as shared control and critical voice. These were not considered to be pertinent to the evaluation of the MGMP Empowerment programme, as the focus was to provide the teachers with the skills and knowledge to help them to move from more traditional practices to ones that were considered to be more student-centred.

2.5.2.9 What Is Happening In this Class? Questionnaire

The What Is Happening In this Class? (WIHIC) questionnaire was developed by Fraser, Fisher and McRobbie (1996) to address contemporary educational concerns. The WIHIC combines modified versions of salient scales from different questionnaires with scales that address concerns such as equity and constructivism.

The WIHIC is available in two versions, namely, a class form (which assesses a student's perceptions of the class as a whole) and a personal form (which assesses a student's personal perceptions of his her role in a classroom). The original 90-item version was later refined to include 56 items in seven scales, these being Student Cohesiveness, Teacher Support, Involvement, Investigation, Task orientation, Cooperation and Equity (Aldridge, Fraser & Huang, 1999; Dorman, 2003). Two example items include 'I give my opinions during class discussions' (Involvement) and 'I receive the same encouragement from the teacher as other students do' (Equity).

The WIHIC is perhaps the most widely used of all of the learning environment questionnaires. It has been translated into numerous languages and used in many countries. Aldridge, Fraser and Huang (1999) cross validated the WIHIC with a sample of 1879 high school students in 50 classes in Taiwan and 1081 high school students in 50 classes in Australia (Aldridge, Fraser & Huang, 1999).

The WIHIC has also been validated in independent studies in Singapore (Chionh & Fraser, 2009; Khoo & Fraser, 2008), India (Koul & Fisher, 2005), South Africa (Aldridge, Fraser & Ntuli, 2009), Indonesia (Fraser, Aldridge & Adolphe, 2010; Wahyudi & Treagust, 2004), Korea (Kim, Fisher & Fraser, 2010), the United States (Allen & Fraser, 2007; Ogbeuhi & Fraser, 2007; Wolf & Fraser, 2008), the United Arab Emirates (Afari, Aldridge, Fraser & Khine, in press; MacLeod & Fraser, 2010), Canada (Zandvleit & Fraser, 2005) and Australia, Canada and the United Kingdom (Dorman, 2003).

Perhaps the most comprehensive validation of the WIHIC was by Dorman (2003) who used a sample for 3980 high school students from Australia, Canada and the UK. His analysis of the data utilised a confirmatory factor analysis that supported the seven scale *a priori* factor structure.

Those studies described in this section supported the validity and reliability of WIHIC. Given the wide use and effectiveness of the WIHIC and its applicability to

secondary school classrooms, it was considered to be most suitable for use in my evaluation of a teacher professional development programme (see Chapter 3, Section 3.5.1.1).

2.5.3 Past Research within the Field of Learning Environments

As discussed in the previous sections, numerous questionnaires have been developed to assess students' perceptions of their classroom learning environments (Fraser, 2007). The present research drew on the learning environment questionnaire described in the previous sections to help to evaluate the effectiveness of the teacher professional development programme. The questionnaire had the potential to provide a range of information about factors such as: whether a class was student-centred or teacher-centred; whether students were active or passive in class; whether students were undertaking collaborative work; and whether the teacher was approachable and supportive (Fraser, 2007, 2012).

Past research within the field of learning environments provided many research models and methods that were of relevance to the present study. Reviews of classroom environment research (Fraser, 2007, 2012) have delineated at least 10 lines of research, including: identifying differences in perceptions of the classroom environment between students and their teachers (Fisher & Fraser, 1983); identifying exemplary teachers (Waldrip, Fisher & Dorman, 2009); and guiding teacher's decisions about implementing strategies to change their classroom environments (Aldridge, Fraser, Bell & Dorman, 2012; Aldridge, Fraser & Sebela, 2004; Waldrip, Reene, Fisher & Dorman, 2008). Determinants of the classroom environment have also been identified, including: cultural differences involving the race of the teacher in Brunei (Khine, 2002); the socio-cultural beliefs of students (Jegede, Fraser & Okebukola, 1994); in Korea, whether students were science or humanities orientated (Lee, Fraser & Fisher, 2003); and gender difference across several countries (Chionh & Fraser, 2009; Quek, Wong & Fraser, 2005). Large cross-national studies have been carried out for the purpose of gaining new insights into

areas such as teaching methods and student attitudes that might be overlooked within one culture (Aldridge, Fraser & Huang, 1999; Aldridge, Fraser, Taylor & Chen, 2000; Fraser, Aldridge & Adolphe, 2010).

Of particular relevance to the present study was past research related to the evaluation of educational innovations. This area of past research is reviewed below.

1.1.4 2.5.4 Evaluation of Educational Innovations

An important application of learning questionnaires in past research has been as a source of process criteria for evaluating the impact of innovations in the classroom (Fraser 2007, 2012). In Singapore, classroom environment instruments have been used to evaluate educational innovations including computer-assisted learning (Teh & Fraser, 1995a, 1995b) and computer courses for adults (Khoo & Fraser, 2008). In Australia, Aldridge and Fraser (2008) used the TROFLEI in monitoring and evaluating the success of an innovative new senior high school on Western Australia in promoting outcomes-focused education. With a sample of 449 students, this study reported changes in student perceptions of the classroom learning environment over the four year programmes.

The use of learning environment criteria has illuminated the impact of a wide range of new educational programs or approaches including computer-assisted learning in Australia (Maor & Fraser, 1996, 2005) and Canada (Raaflaub & Fraser, 2002); innovations involving anthropometry activities in science education in the United States (Lightburn & Fraser, 2007); Year 11 earth science in Korea (Cho, Yager, Park & Seo, 2004); inquiry-based science instruction for middle-school students (Wolf & Fraser, 2008); an innovative science course for prospective elementary students (Martin-Dunlop & Fraser, 2008); and the effectiveness of the Science and Mathematics Integrated with Literary Experiences (SMILE) project carried out with fifth grade students in the United States (Mink & Fraser, 2005).

Nix, Fraser and Ledbetter (2005) evaluated an innovative science teacher professional development programme that was based on the Integrated Science Learning Environment Model. They used the Constructivist Learning Environment Survey (CLES), with a unique side-by-side format, to examine 445 students' perceptions of the learning environment created by the teachers who had attended the professional development opportunity and to compare them with the classes of other teachers who had not. Students perceived the classes of teachers who had attended the teacher development course to have higher levels of Personal Relevance and Uncertainty (as assessed by the CLES) than the comparison classes.

Lightburn and Fraser (2007) used the Science Laboratory Environment Inventory (SLEI) to evaluate the effectiveness of using anthropometry activities in science education. Their study involved a sample of 761 high school students in South Florida and the results indicated that, relative to a comparison group, students' perceptions were more positive on some scales of the SLEI.

Martin-Dunlop and Fraser (2008) used scales from the SLEI and the WIHIC in their evaluation of the effectiveness of an innovative science course for prospective teachers. Using a pre-post design, with a sample involving 525 university students in 27 classes in California, students reported statistically significant improvements on all seven scales assessing the laboratory learning environment and attitudes towards science. The largest gains were observed for Open-Endedness and Material Environment (with effect sizes of 6.74 and 3.82 standard deviations, respectively).

Pickett and Fraser (2009) argued that the litmus test of the success of any teacher professional development programme is the extent of changes in teaching behaviours and ultimately student outcomes in the participating teachers' school classrooms. Consequently, their evaluation of a two-year mentoring programme in science for beginning elementary school teachers drew on the field of learning environments in gauging this programme's success in terms of participants' classroom teaching behaviour as assessed by their school students' perceptions of their classroom learning environment.

The present study drew on and extended research in the field of learning environments by using a learning environment instrument (developed for the purpose of this study) to help to assess the effectiveness of a teacher professional development programme in Indonesia.

2.6 Student Attitudes

In the present study, pre–post changes in student attitudes were used an indication of the effectiveness of the professional development programme. Attitude has been defined and measured in many ways. Attitude is a non-observable psychological process whose presence can only be assumed. Attitudes cannot be observed or measured directly, but rather, their existence must be inferred from their consequences (Mueller, 1986). Thurstone (1928) defined attitudes as the sum total of a man's inclination and feelings, prejudices and bias, preconceived notions, ideas, fears, threats and convictions about a specific topic. The idea that attitudinal behavior is learned and could be modified is widely accepted today by social scientists. According to Cotterall (1995, p. 195), a student's attitude to learning can influence his or her learning behaviour and learning outcomes.

Reid's (2006) definition of attitudes involves three components, namely, cognitive (knowledge of the object, belief or ideas), affective (feelings regarding the object, such as like or dislike) and behavioural (the tendency towards an action or objective). Other researchers have viewed these three components more independently, for example, when we judge something emotively such as good or bad, like or dislike (Crano & Prislin, 2006). Such a definition enables researchers to distinguish attitudes from emotions or behaviours. Kind, Jones and Barmby (2007, p. 873) provide a definition based on these components of attitudes as "the feelings that a person has about an object (evaluative attitudes are always towards something often called an attitude object) based on their beliefs about that object".

Some researchers have emphasised that attitudes involve both cognitive and affective aspects (Bagozzi & Burnkrant, 1979; Oppenheim, 1992; Hilgard, 1980;

Triandis, 1971). Researchers have also indicated that attitudes involve three components: cognition, affect, and conation (the behavioural component). They suggested that feelings toward an attitude object and beliefs about the characteristics of the attitude object impact on behaviour. However, when the use of the term is restricted to only the affective dimension, attitude is regarded as a unidimensional construct (Fazio & Zanna, 1981; Schopler & Insko, 1992).

There have been many types of attitude measures developed over the years. Osborne (2003) reviewed five main methods of assessing attitudes, these being: preference ranking; attitude scales; interest inventories; subject enrolment; and qualitative methods. The use of attitude scales is probably the most common method of assessing attitude. These scales have included a range of response formats and indicators such as the Likert scale, Thurstone scale, Guttman scale and the semantic differential technique.

In the past, research involving students' outcomes has focused primarily on educational objectives in the cognitive domain (Weinburgh, 1995) but, in more recent times, it has been acknowledged that stimulating students' attitudes and motivation to learn is one of the greatest challenges for teachers (Theobald, 2006). Research has revealed that students' attitudes are an important affective component because they play a central role in their conceptual change processes, critical thinking, learning strategies and achievement (Kuyper, van der Werf & Lubbers, 2000; Shulman & Tamir, 1972). Given the importance of student attitudes to a range of student outcomes, the assessment of students' enjoyment of their English classes was included as an important component to help to evaluate the effectiveness of teacher professional development programme in the present study.

In my study, to assess students' attitudes in English lessons, I modified an attitude scale based on the Test Of Science-Related Attitudes (TOSRA, Fraser, 1981). The TOSRA was selected for use in this study because it overcomes most of the problems addressed associated with past questionnaires developed to assess students' attitudes (Kind, Jones & Barmby, 2007). The scales of the original TOSRA

are based on Klopfer's (1976) classification of students' attitudinal aims: attitude to science and scientists, attitude to inquiry, adoption of scientific attitudes, enjoyment of science learning experiences, interest in science, and interest in a career in science.

One of the scales included in the TOSRA, the Enjoyment of Science Lessons, was selected and modified for use in the present study. This scale has been validated and used in numerous past studies (Martin-Dunlop & Fraser, 2008), in different countries such as Australia (Fraser & Butts, 1982; Fraser & Fisher, 1983), Singapore (Wong & Fraser 1997), Korea (Lee, 2001), Indonesia (Aldridge et al., 2010) and the United States (Farenga & Joyce, 1998; MacDowell-Goggin & Fraser, 2004; Martin-Dunlop & Fraser, 2008; Soto-Rodriguez & Fraser, 2004). The scale also has been modified for use in different subject, including mathematics (Taylor, 2000) and geography (Walker, 2006). In addition, the TOSRA has been modified and used successfully in a variety of settings and for different courses. This scale was selected for its usefulness in terms of assessing students' enjoyment of a subject. The version selected for use in the current study was adapted by Aldridge, Fraser and Huang (1999) was been found to be reliable in past studies (Aldridge & Fraser, 2008; Aldridge, Fraser & Huang, 1999). The modification and translation of the enjoyment scale to suit English classes in the Indonesian setting is described in Chapter 3.

2.7 Chapter Summary

The overarching aim of my study was to investigate the effectiveness of a teacher professional development programme. Thus, my review of literature aimed to present a succinct and coherent picture of teacher professional development and a background to two important aspects of the evaluation: the field of learning environment and students attitudes.

Although teacher professional development has been defined in many different ways, my study drew on Hassel's (1999) definition, which describes professional development as the process of improving staff skills and competencies needed to

produce outstanding educational results for students. Despite the differences in definitions, there appears to be some consensus with respect to categorising professional development into four overarching paradigms: the deficit paradigm; the professional growth paradigm; the educational change paradigm; and the problem-solving paradigm. In the past, much professional development has centred on more traditional paradigms, however there is currently a move away from these approaches (in which teachers are passive recipients) towards an emphasis on to more complex processes that involve teachers playing an active role in shaping their professional growth through reflective practice.

Effective teacher professional development is influenced by a multitude of factors. The three most influential characteristics are the content of the professional development (the knowledge skills and understanding), the process variables (the type, form and planning of activities) and the context variables (the organisation, system and culture in which the professional development takes place). Research evidence suggests that there are a number of factors that are most likely to influence the success of professional development and these were used by Fishman et al (2003) to develop a model of professional development.

To date, much professional development is not evaluated in a systematic way, however there is a growing interest in doing so. A number of models have been developed to guide the evaluation of professional development. This chapter reviews the three models, developed by Mathison (1992), Gusky (2000) and Fishman et al (2003), that were drawn on to develop the model of evaluation used in this study.

This study drew on and extended the field of learning environments as part of the evaluation of the professional development programme. The field of learning environments has grown over the past few decades following the pioneering work of: Lewin (1936) who connected personality and environmental influences to behaviour; Murray (1938) who expanded Lewin's work accounting for personal needs, environmental presses and differences perceived by observers and

participants; and Stern, Stein and Bloom (1956), who elaborated further on Murray's work in identifying that individuals might have personal perceptions that differ from their perceptions within a group. Human interactions have been categorised by Moos (1974) into three dimensions of relationship, personal development and system maintenance and change, which have served as the framework for assessing various aspects of learning environments.

During the growth of this field, numerous instruments have been developed, modified and used to assess dimensions of the learning environments at all levels, from primary school through to university. The instruments have been used in various subject areas including science, mathematics and computing. Many of these instruments have been found to be reliable and valid in a range of countries and contexts, including the Constructivist Learning Environment Survey (CLES), What Is Happening In this Class? (WIHIC) and Science Laboratory Environment Inventory (SLEI).

The What Is Happening In this Class? (WIHIC) was selected for use in this study because of its highly robust nature, effective use in numerous settings around the world and its personal form that can be used to assess sub-groups (such as students in urban and rural settings). This chapter reviewed the development and validation of the WIHIC in different settings.

These learning environment instruments have been used in numerous lines of learning environment research. This chapter reviews past research in the field of learning environment, with particular attention to reviewing literature on that past research related to the use of learning environment dimensions as criteria in the evaluation of educational innovations.

Students' attitudes were assessed in the current study using a modified version of the Test of Science Related Attitudes (TOSRA) because of its effective use in numerous past studies. Therefore, this chapter reviews literature related to the measurement of student attitudes.

The following chapter outlines the research methods used in the present study, which examined the effectiveness of a professional development programme.

Chapter 3

RESEARCH METHODS

3.1 Introduction

The previous chapter provided a review of literature related to the present study. The review highlights that, despite the importance of professional development in the era of educational reform, there are serious concerns about the evaluation of the effectiveness of professional development practice (Guskey, 2000). Additionally, it was noted in Chapter 2 that there is a greater need for a more robust system for evaluation of teacher professional development activities (Guskey, 2000; Sparks & Hirsch, 2002).

This chapter describes the research methods used in the present study under the following headings:

- A New Model for Evaluating Professional Development (Section 3.2);
- Research Design (Section 3.3);
- Sample (Section 3.4);
- Data Collection Methods (Section 3.5);
- Data Analysis (Section 3.6);
- Ethical Considerations (Section 3.7); and
- Chapter Summary (Section 3.8).

3.2 A New Model for Evaluating Professional Development

An important component of any professional development programme is the translation of what has been learned into classroom practice (Fullan, 2003; Guskey, 2000). Hitherto, however, much evaluation related to the effectiveness of

professional development programmes has either omitted this aspect of the evaluation altogether or not taken into consideration the impact of the evaluation on changes in classroom practice and student outcomes (Guskey, 2000).

My review of literature indicated that, of the available models for evaluating professional development, a composite model, drawing on salient features of existing models would be most suitable to evaluate the MGMP Empowerment programme. Many existing models of professional development evaluation have focused on the participants' perceptions of the programme experience, student performance, context variables, impacts of delivery models and student learning. However, these evaluation models rely on feedback from the teachers as the participants of professional development and, sometimes, a non-participant observer as the evaluator.

My review of literature suggests that the inclusion of student perceptual measures might provide a valuable indicator of effectiveness of the MGMP Empowerment programme. It was considered that, because students spend most of their time in the classroom, their reactions to and perceptions of their learning experiences were significant. With this in mind, I developed a model of evaluation which was influenced by three models developed by Mathison (1992), Guskey (2000), and Fishman, Mark, Best and Tal (2003) and incorporated an aspect that would examine the students' perceptions of their learning experiences. Each of the three models emphasised the impact of the professional development in terms of pedagogical changes in classroom and students' outcomes (see Section 2.4). By including the salient aspects of these three existing models and incorporating student perceptual measures into my new model, I was able to provide a holistic evaluation that suited to the Indonesian context.

My study endeavoured to overcome past problems, related to the evaluation of teacher professional development, by modifying and combining components of three existing models. Importantly, the present study also drew on and extended past learning environments studies by incorporating students' perceptions as an

integral component of the evaluation model to examine whether their perceptions of their learning experiences and attitudes changed before and after the professional development programme had been introduced.

In adapting the three models, it was necessary to change the sequence of Guskey's (2000) five levels of evaluation to successfully combine them with Mathison (1992) and Fishman et al.'s (2003) models. In some cases, elements from all three models were similar and these three were brought together to assess a similar aspect. My modified model involved five phases. These phases are presented diagrammatically in Figure 3.1 and are described below.

- The first phase involved the collection of important baseline data which examined both students' perceptions of their learning experiences and the teachers' classroom practices and teaching strategies prior to the commencement of the professional development.

- The second phase involved evaluating teachers' views of the relevance and utility of the programme in terms of their views towards the professional development experience and whether the programme had expanded their knowledge and skills.

- The third phase involved examining the extent to which the ideas and concepts, imparted during the professional development programme, were translated into practice by the teachers.

- The fourth phase examined whether students perceived the classroom learning environment differently and whether their attitudes toward English classes had changed.

- The fifth phase of the model involved examining contextual factors that might promote or impede the translation of ideas imparted during the professional development programme.

Thus, my evaluation model provided a framework that was used to guide the gathering of data at each phase of the evaluation. Figure 3.1 portrays these five phases and how they interrelate.

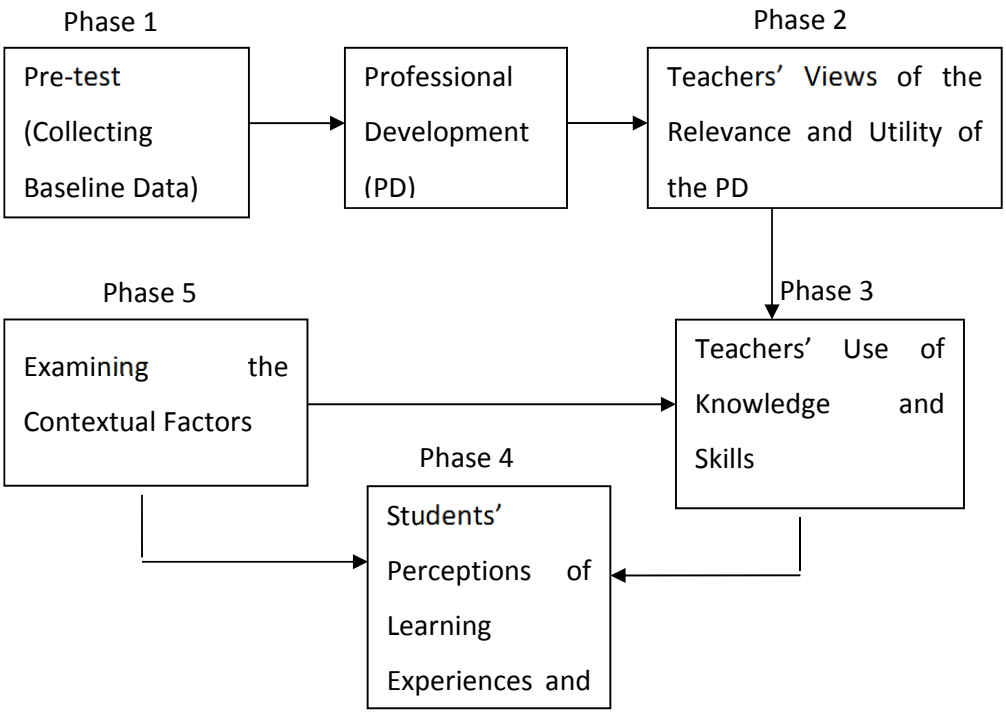


Figure 3.1 A New Model of Teacher Professional Development Evaluation

The next section describes the design of the research, based on the evaluation model and the methodology employed.

3.3 Research Design

The evaluation of the impact of the MGMP Empowerment professional development programme involved a mixed-methods approach in which, multiple research methods were drawn on and combined. The mixed-methods approach involved collecting, analysing and integrating both quantitative and qualitative data in a sustained programme of inquiry (Creswell & Plano Clark, 2010).

The evaluation of students' perceptions of the learning environments and evaluation of student outcomes (Phase 1 and Phase 4), involved a pre–post design in which quantitative data was collected prior to the start of the professional development programme and then again at the end of the professional development programme. This component of my study was conceptualised as the relationship between 'focal' and 'reference' variables design (William, 1998). The 'focal' variable consisted of the student perceptions of the classroom psychosocial environment and their Enjoyment of English classes, while the 'reference' variable was the on-going teacher professional development provided as a part of the MGMP Empowerment programme. This stage of the study employed a more positivistic framework, favouring an objectivist view.

The evaluation of the teachers' attitudes and views (Phase 2), their use of knowledge and skills (Phase 3) and the examination of contextual factors that could promote or impede the implementation (Phase 5) involved the collection of qualitative data and employed an interpretative framework drawing on elements of constructivism (Schwandt, 1994; Tobin, 1993). The collection of qualitative data added richness and depth to the evaluation as a whole. During these phases, the focus was on understanding and interpreting phenomena and making meaning out of the quantitative data.

The qualitative component of the evaluation employed a case study approach (Stake, 1995; Yin, 2011) involving the collection of data through teacher reflective journals, classroom observations and in-depth interviews with teachers and students. This information helped to explain, clarify and contextualise the quantitative overview, whilst also provided causal information for the quantitative results, thereby building on the strengths of both the quantitative and qualitative research methods (Creswell & Plano Clark, 2010; Creswell, Plano Clark, Gutmann & Hanson, 2003). Thus, embracing a mixed-method approach for this study opened the door for different worldviews, forms of data collection and analyses (Creswell & Plano Clarke, 2010).

For each phase of the evaluation, the following sections describe the samples involved (Section 3.4) and the methods used (Section 3.5).

3.4 Sample

The selection of the sample was an important factor for each phase of the evaluation. In the social and behavioural sciences, there are two groups of sampling procedures: probabilistic sampling; and purposeful sampling (Creswell & Plano Clark, 2010). Probabilistic sampling is intended to select a large number of individuals who are representative of the population and purposeful sampling involves the researcher intentionally selecting teachers who have experienced the central phenomenon or the key concept being explored. In my study, sampling involved both of these procedures to recruit teachers and students to evaluate the impact of the MGMP Empowerment programme.

To select the sample for the quantitative overview (used in Phases 1 and 4 of the evaluation), multiple cluster sampling was employed. According to Teddlie and Yu (2007), multiple cluster sampling involves implementing two basic probability sampling techniques, in conjunction with one another, to generate a more complex sample. The first stage of sampling involved the random selection of an initial cluster (in this case I selected the region or district from which the English teachers came). The second stage of sampling involved the random selection of a second cluster (English classes).

Based on this sampling, six out of 35 districts from within the Central Java Province were randomly selected. The 33 participating teachers and their 32 schools from these selected districts were then approached to determine whether they would be willing to participate in this component of the study. During this stage, the 32 schools, selected for inclusion, were categorised according to school type (public or private) using the schools' profiles.

The schools from which the 33 teachers were selected represented a range of classroom sizes (the smallest having approximately 21 students and the largest having 46 students). The schools were located in urban (*n*=15) and rural (*n*=17) areas of Central Java, Indonesia. To distinguish between urban and rural schools I drew on the United Nation's (2010) definition, which takes into account the geographic location of the school, the size of population surrounding the school (that is, urban 10,000 – 24,999; rural < 10,000) and the distance of the school from the service centre (hospital, shopping centre, public service central office). Given the wide range of schools from different regions, this sample was considered to be representative of schools in these areas.

Table 3.1 provides a breakdown of the sample of the 33 teachers who were involved in my study. Of the teachers involved in the study, 16 were teaching in urban schools and 17 were teaching rural schools. The sample involved 16 male teachers and 17 female teachers. Their teaching experience varied from ten to 15 years (10 teachers), 16 to 20 years (15 teachers) and more than 20 years (8 teachers).

Table 3.1 Breakdown of the Teacher Sample by Locality, Gender and Teaching Experience

	Gender		Teaching Experience (in Years)		
School Type	Male	Female	10 to 15	16 to 20	20 up
Urban	7	9	6	7	3
Rural	9	8	4	8	5
Total	16	17	10	15	8

To collect the data for this pre–post component of the evaluation (Phases 1 and 4), I worked with the 33 participating teachers, to randomly select two classes for each teacher that would be involved (by drawing numbers representing the classrooms that they taught out of a box). This provided a sample of 66 classes, 32 from urban schools and 34 from rural schools. The sample included only those students who were present for both the pre-test and the post-test. The student sample, reported

in Table 3.2, involved a total of 2,417 students (912 boys and 1,505 girls) in grades 7 (*n*= 412), 8 (*n*= 585) and 9 (*n*= 1,420). This large sample size was considered to be advantageous as the results were likely to be more generalisable.

Table 3.2 Breakdown of the Student Sample by Gender and Locality

Level	Number of Students by Gender		Number of Students by Locality		Total
	Male	Female	Urban	Rural	
7	184	228	184	228	412
8	277	308	354	231	585
9	451	969	422	998	1420
Total	912	1505	960	1457	2417

Qualitative information was gathered at various intervals during the one-year programme. A case study approach was involved for each of the phases of the evaluation. A sample of six case study teachers were selected from the 33 teachers described above. The selection of case study teachers for the present study involved non-probability purposive sampling (Teddlie & Yu, 2007) to identify two categories of schools within the population that met specific criteria. The criteria for selection included schools that were different from each other in terms of geographical location (i.e. rural and urban). The decision to include teachers from schools that were different from each other was made to provide data that was more varied, thereby increasing the likelihood of being able transfer the results to other contexts or settings. Of the six case study teachers, three teachers were teaching in rural schools and the other three were teaching in urban schools.

To provide insights into teachers' views towards the various activities presented during the professional development programme (Phase 2 of the evaluation model), all of the participating English teachers (*N*=138) were involved. The teachers all wrote reflective journals during the teacher professional development programme (detailed in Section 3.5.2.1). At the end of the programme, the reflective journals of all teachers (*N*=138) were analysed until saturation was reached (e.g. no new

information was found). Although data saturation occurred after first 30 reflective journals, all reflective journals were read and all content was coded (Clayton and Thorne, 2000). Analyses suggested a high level of homogeneity among the population and, as recommended by Guest, Bunce and Johnson (2006), it was decided that the reflective journal of the six case study teachers were sufficiently representative of the whole sample.

To evaluate whether the teachers believed that the professional development programme had expanded their knowledge and skills (Phase 2), a sample of 33 participating teachers (who were randomly selected from the 138 teachers, at the first stage of sampling for my study) were interviewed. The 33 teachers whom were interviewed taught different grade levels (grades 7, 8 and 9) and represented the geographic, community size, and school size distributions of Central Java province. Most of the teachers were between 30 and 45 years of age, with a teaching experience of between 10 and 25 years. Interview data saturation occurred after 20 interviews, as suggested by Kvale (1996) and Guest et al. (2006). Analyses suggested a high level of homogeneity among the sample and, it was decided that the interview data of the six case study teachers were, again, sufficiently representative of the whole sample (Guest, Bunce & Johnson, 2006).

To assess changes in teachers' practices (Phases 1 and 3), the six case study teachers' classrooms were observed for minimum 4 hours (one observation was conducted before the commencement of professional development programme – Phase 1 – and the rest were conducted by the end of the programme – Phase 3). The classroom observations were used to provide information about how the teachers used the knowledge and skills that they gained during the professional development programme in their day-to-day practices.

The fourth phase of evaluation involved examining changes in students' experience and their attitudes towards English as indicators of the impact of teacher professional development programme. For this phase, both quantitative and qualitative data were employed (sample for the quantitative aspect was explained

previously). For the qualitative component, a total of twelve students (two students from one class of each case study teacher) were selected by implementing a purposive sampling technique (Tashakkori & Teddlie, 2003). Although the sample was based largely on the students' willingness to be involved, wherever possible, care was taken to provide a sample that was representative of gender and academic abilities in the classroom.

The fifth phase of evaluation involved the evaluation of contextual issues impeded the implementation. For this purpose, the six case study teachers were interviewed to examine whether contextual factors existed that might promote or hinder the implementation of ideas presented during the professional development programme.

3.5 Data Collection Methods

Whereas the previous section described the sample for the various phases of the evaluation, this section describes the data collection methods used for both the quantitative (Section 3.5.1) and qualitative (Section 3.5.2) data.

1.1.5 3.5.1 Quantitative Data Collection

Quantitative data was collected in Phases 1 and 4 of the evaluation to examine whether changes occurred, from the students' perspective, at the classroom level. In these two phases of the evaluation model, a pre–post design was used to: 1) examine the extent to which students' perceived changes in their learning experience in terms of changes the learning environment and 2) investigate whether students' enjoyment of their English lessons changed. This involved administering two instruments before the professional development started (Phase 1) and one year later at the end of the programme (Phase 4). This section describes the selection and modification of the surveys to examine: changes in student learning experiences (Section 3.5.1.1); and changes in students' attitudes (Section 3.5.1.2). The section goes on to describe the translation of the surveys into

Indonesian (Section 3.5.1.3), the pilot study (Section 3.5.1.4) and the procedures used to administer the surveys (Section 3.5.1.5).

3.5.1.1 Assessing Changes in Students' Learning Experiences

Although past research has relied on students' cognitive outcomes to evaluate the effectiveness of teacher professional development, these measures do not give a complete picture of teachers' behavioural changes in the classroom level. To examine the extent to which teachers change their pedagogical practice, the present study drew on and extended past research in the field of learning environments. Learning environment refers to the tone, ambience or atmosphere created by a teacher through the relationships developed within the classroom and the way in which instruction is delivered (Fraser, 2012). Because students and teachers spend many hours in classrooms, they have a large stake in what happens to them and their reactions and perceptions of their classroom experiences are significant. Therefore, the personal nature of the perceptions of those who are in the classrooms on a daily basis can provide a wealth of information into what happens at this level.

A review of literature indicated that the What Is Happening In this Class? (WIHIC) was a suitable instrument to help to evaluate students' perceptions of their learning experiences as a means of gauging the translation of what was imparted during the professional development programme into classroom practice. The WIHIC was originally developed by Fraser, Fisher and McRobbie (1996) to assess students' views of a contemporary, student-centred learning environment. The scales of the WIHIC are underpinned by the constructivist theory of learning and have been found to be strong predictors of student outcomes (Aldridge, Fraser & Huang, 1999; Fraser, 2012). The dimensions assessed in the WIHIC provide information, from the students' perspective, about how learning opportunities are structured and provide an indication of the emphasis that teachers place on student learning (Aldridge, Fraser & Huang, 1999). For example, the Involvement scale measures the extent to which students have attentive interest and participate in discussion. I considered,

therefore, that the WIHIC scales would be valuable indicators of changes in teachers' practices over the course of the professional development programme.

A review of past research related to the field of learning environment suggested that What Is Happening In this Class? (WIHIC) questionnaire is one of the most widely-used learning instruments around the world (Fraser, 2007). Chapter 2 provides an overview of the countries, subjects and grade levels, in which the WIHIC has been used and found to be reliable. As the review outlines, the WIHIC has also been used and validated in Indonesia (Fraser, Aldridge & Aldophe, 2010; Margianti, Aldridge & Fraser, 2004). However, these studies in Indonesia were conducted either in science classes or at the tertiary-level in mathematics classes. Given that the present study sought to evaluate a teacher professional development programme aimed at improving the teaching strategies of high school English teachers, it was necessary to modify the WIHIC to ensure the suitability of the instrument.

As a first step, the scales of the WIHIC were inspected to ensure their relevance to English language classes. At this stage, the Investigation scale (originally developed to assess students' perceptions of emphasis on the skill in problem solving and investigations in science) was not considered to be relevant and was, therefore, omitted. All of the remaining scales were considered to be relevant.

As a next step I scrutinised individual items to ensure that they were relevant to the Indonesian context and to English language classes. All of the items were considered to be relevant and were used in the Indonesian version of the WIHIC. To make sure that students were clear that their responses were related to his/her English teacher, the word 'English' was added where appropriate. For example, "The teacher takes personal interest in me" was reworded to "The English teacher takes personal interest in me". Finally, the scales of the WIHIC were examined to ensure that all constructs, considered important to English language classes in Indonesia, were included.

As an important element of English language learning involves providing students with opportunities to develop skills related to locating relevant information, I developed a scale to assess the extent to which students perceived that they were required to find information (such as checking the dictionary or using a reference book). The addition of this scale, which I named Finding Reference scale, was considered particularly pertinent given that the content of the professional development programme included examining and using teaching materials to help teachers to design such tasks.

The modified WIHIC questionnaire, used in my study, consisted of 56 items in seven scales, namely, Student Cohesiveness, Teacher Support, Involvement, Finding Reference, Task Orientation, Cooperation and Equity. Table 3.3 provides a scale description and sample of item for each of the scales used in my study. A copy of the English version of the modified WIHIC used in my study is provided in Appendix C.

Table 3.3 Description and Example Item for Each WIHIC Scale

Scale Name	Scale Description	Sample Item
	The extent to which…..	
Student Cohesiveness	students know, help, and are supportive of one another.	I help other class members who are having trouble with their work.
Teacher Support	the teacher helps, befriend, trust, and shows interest in students.	The teacher considers my feeling.
Involvement	students have attentive interest, participate in discussion, perform additional work and enjoy the class.	I give my opinion during class discussions.
Finding Reference	There is an emphasis on the skills and processes required to use references to solve the problems.	I solve problems by using information obtained from references.
Task Orientation	it is important to complete activities planned and to stay on the subject matter.	I am ready to start this class on time.
Cooperation	students cooperate rather than compete with one another on learning tasks.	I cooperate with other students when doing assignment work.
Equity	students are treated equally by the teacher.	I receive the same encouragement from the teacher as other students do.

3.5.1.2 Assessing Changes in Students' Attitudes

To assess students' attitudes towards their English class, I used a modified scale from the Test of Science-Related Attitudes (TOSRA, Fraser 1981). The widely-used TOSRA was selected because it has been modified and validated in a range of studies. Although originally developed for use in science subjects, the attitude scale has been adapted for use in a range of subjects including mathematics (Teng & Wong, 2009, Dorman, et.al., 2002), computing (Khoo & Fraser, 2008, Zandvliet & Fraser, 2005; Soerjaningsih, Fraser & Aldridge, 2010) and geography (Walker, 2006). The strong reliability and wide applicability in different countries and subjects made this attitude scale a suitable choice for the present study. For the purpose of this study, one of the scales, the students' enjoyment of science classes was modified to assess students' enjoyment of their English classes (in terms of whether they viewed them as interesting, boring, dull or exciting).

To ensure the suitability of the scale, several modifications were made to the items. First, the wording of all items was changed from 'Science Lessons' into 'English Lessons'. For example, "I look forward to science lessons" was reworded to "I look forward to English lessons". Second, to avoid the students' confusion, items that were originally worded in a negative manner such as "I dislike science class" were rephrased in a positive manner, such as "I like English lessons". The modified Enjoyment of English classes scale consisted of 10 items.

To minimise confusion and to provide consistency in the response format, I changed the Likert-type response format of the original scale (Strongly Agree, Agree, Not Sure, Disagree, and Strongly Disagree) to the frequency response format used in the WIHIC. In this way, both instruments used a five-point frequency response scale of Almost Never, Seldom, Sometimes, Often and Almost Always. This change also required some changes to wording to individual items to ensure that they were suitable for use with the new response format. A copy of the English version of the Enjoyment of English Classes scale is provided in Appendix D.

3.5.1.3 Translation of the Surveys

Once both of the instruments had been modified to suit the Indonesian context and, specifically, for use in English classrooms, the next step was to translate them into Indonesian. The translation of the two questionnaires into Bahasa Indonesia involved the rigorous process of back translation (Bracken & Barona, 1991; Brislin, 1970, 1980; Chapman & Carter, 1979). The first step of the process involved the researcher translating the instrument into the Indonesian language. Second, an independent translator conversant in both English and Indonesian, but who had not seen the original questionnaires, translated the items back into English. The back translation was then used to compare the two versions to ensure that the Indonesian version maintained the meanings and the concepts included in the original version. The process was reiterated until both original and back-translated versions of two questionnaires (WIHIC and TOSRA) were considered to have achieved semantic equivalence. A copy of the Indonesian version of the WIHIC and attitude scales is provided in Appendices C and D, respectively.

3.5.1.4 Pilot Study

Prior to the main data collection, a pilot study involving 89 students from grades 7, 8 and 9, who were not involved in the evaluation component of the study was conducted. The pilot study provided an opportunity to simulate the main study to examine:

1) the comprehensibility and clarity of the items in the Indonesian version of WIHIC and attitude scale;
2) the response format of the WIHIC and attitude scale;
3) the procedures for data collection; and
4) the amount of time required by students to complete each of the instruments.

During the pilot study, the 89 students were asked to complete both questionnaires (WIHIC and attitude scale) and six of them were interviewed after the completion. The insights emerging from the pilot study are summarised below:

1) The comprehensibility and clarity of the items in the Indonesian version of WIHIC and Attitude Scale was investigated. Students' responses to the surveys and interviews with the students indicated that the wording of items were clear. It was decided that the surveys were comprehensible to be used in the main study.

2) Students were able to effectively use the response format. They found the instruments to be clear and did not indicate any difficulties in responding. It was decided, therefore, that the format for the questionnaires was suitable for the main study.

3) The procedures used to administer the survey in the pilot study proved workable. The directions were considered to be straightforward with the students following step-by-step.

4) The amount of time required by students to complete the surveys was approximately 10 minutes, which was within expectations.

3.5.1.5 Administration of the Surveys

The collection of data using the two instruments was conducted in two phases, these being: the pre-test (before the start of the professional development programme); and post-test (after one year at the end of the programme). The pre-test was administered after the students had been in classes for eight weeks to provide sufficient time for a stable learning environment to be established and to ensure that students were familiar with their teachers (as recommended by Wubbels & Brekelmans, 2005). To examine whether the students' perceptions of learning environment and attitudes changed, the same surveys were administered towards the end of the school year, at which time the MGMP Empowerment programme was finished.

To increase the likelihood of students being comfortable when responding to the items, I personally administered the surveys and the teachers were asked to stay

out of the classroom during this time. Prior to administration, I gave directions to the students in their language and clarified any questions that they had. At this time, I also advised the students that their responses were confidential and would not be seen by their teachers.

After completion of the questionnaires, the data were entered into an Excel spread sheet, by hand, in preparation for analysis. Incomplete questionnaires were discarded and, in cases where an error or missing response occurred, a '3' was recorded as a mid-range indicator. The questionnaire data only included responses for those students who were present for both the pre-test and the post-test. 15 of the surveys that were responded to by students (six for the pre-test and nine for the post-test) were considered to be unusable and, therefore, discarded.

The following sections describe the methods used to collect the qualitative data.

3.5.2 Qualitative Information Gathering

Phases 1, 2, 3 and 5 of the model all involved the collection of qualitative data. Phase 1 involved one-hour of observations before the professional development started. Phase 2 involved examining teachers' views towards the teacher professional development experience and whether they believed that the programme had expanded their knowledge and skills. Phase 3 involved evaluating the extent to which teachers used new knowledge and skills in their classroom practice and Phase 5 involved examining contextual factors that promoted or impeded the implementation of ideas. The qualitative information were gathered at various intervals during the one-year academic programme. This section describes the methods used to gather qualitative information, including: teacher reflective journals (Section 3.5.2.1), classroom observations (Section 3.5.2.2) and in-depth interviews (Section 3.5.3.3).

3.5.2.1 Teacher Reflective Journals

The second phase of the model involved the evaluation of the teachers' reactions to the professional development experience and whether they felt that they had learned new knowledge and skills. Teacher reflective journal were used to provide information to illuminate this aspect. Prior to data collection, I requested permission from all of the teachers (*N*= 138) to use their reflective journal as data for evaluation. All of the teachers agreed. Teachers were provided with structured time, during the professional development programme, to reflect on the experiences provided and to record their feelings in the journal. The teachers were given guiding questions to help them to record and structure their reflective journals. The guiding questions, used to help structure the journals, were adapted from Guskey's (2000) 'Response Evaluation Form'. Appendix E provides a copy of Response Evaluation – Teacher's Reflective Journal. The teachers were given 10 to 15 minutes at the end of each day, during the professional development programme, to record their reflections in their journals. Structured time for reflection was specified to assure some level of consistency across all teachers.

3.5.2.2 Classroom Observations

Phases 1 and 3 of the evaluation model involved the examination of whether teachers made changes to their classroom practices that were consistent with the pedagogical ideas introduced during the professional development programme. Classroom observations were conducted in the classes of the six case-study teachers (the selection of whom is described in Section 3.4). A minimum of four hours of observations were made in each of the classes of the case study teachers during the school year. At Phase 1, the classrooms of each case study teachers were observed for one hour prior to the commencement of the teacher professional development programme. These observations were used to gather information about the teaching strategies that the case study teachers used. The observations for the Phase 3 were conducted near the end of the professional development programme. These classroom observations focused on assessing the extent to which the teachers utilised the suggestions proposed at the teacher professional

development programme and to determine the effectiveness of these implementations at the classroom level.

My position during the classroom observations was as a non-participant observer and I seated myself at the back of the class to be less obtrusive. Observations were recorded using a classroom observation checklist. The observation checklist was used to help me to assess the extent to which teachers utilised the suggestions proposed at the teacher development workshops in a way that was consistent across all observations. A copy of the observation checklist can be found in Appendix F. All observations were video recorded for later analysis. After each observation, I discussed my notes and checklist with the teacher whom I observed, providing them with an opportunity to provide feedback. This member checking helped to ensure that my interpretations were accurate to ensure that the observation data was valid.

3.5.2.3 Interviews

As part of the second and fifth phases of the evaluation model (see Figure 3.1), in-depth, semi-structured interviews were conducted with each of the six case study teachers (Kvale & Brinkmann, 2009). In-depth interviews were used to investigate: 1) how they reacted to the professional development experiences (attitudes) (Phase 2 of the evaluation model); 2) their views about whether they had expanded their knowledge and skills a result of participating in the programme (Phase 2 of the evaluation model); and 3) contextual factors viewed as promoting or impeding the process of implementation (Phase 5 of the evaluation model).

In-depth interviews were selected for two reasons. Firstly, it provided an ideal means of exploring the teachers' views about various aspects of the teacher professional development programme and its implementation in their classrooms. Secondly, in-depth interviews helped to build a good relationship with the teachers. Such a relationship helped to elicit more detailed information about the teachers' personal beliefs and theories of language learning and teaching; details that might not have been possible to access through observations alone (Denzin & Lincoln,

2005). In-depth interviews were considered to be advantageous over other types of data collection methods because of their interactive nature which allowed me to probe more deeply in a bid to understand the views of the participants. In this study, interviews were conducted with each of the six case study teachers and two of each of their students, both of which are described below.

Semi-structured interviews were used to provide a degree of consistency across interviews. This format provided a framework that allowed flexibility and a more conversational style, two-way dialogue. Semi-structured interviews were preceded by each of the observations, to provide me with a clearer understanding of the observations that were to follow and again after the observations to discuss the issues about the teaching strategies observed during the observations (Kvale & Brinkmann, 2009).

Teacher Interviews

In the present study, the semi-structured interviews provided me with opportunity to obtain feedback and to explore how the teachers responded to the professional development programme as well as their views on the programme's effectiveness. The interviews with case study teachers also helped to direct the formulation of tentative assertions and to inform subsequent observations. Interviews were held with each of the case study teachers at the beginning and end of each observation visit.

The semi-structured interviews ranged from informal to formal. Formal interviews with each of the six case study teachers were conducted after the classroom observations. In addition to the semi-structured interviews that were conducted formally, potential informal conversations were also held prior to classroom observations (Kvale & Brinkmann, 2009). These provided valuable opportunities to discuss the school conditions as well as to get to know the teachers better personally.

To ensure the suitability of the questions, the interview schedules used for the study were piloted with two teachers who were not involved in this study. All of the interviews were audio recorded and later were translated, paraphrased and transcribed for analysis. Appendix I provides an example interview schedule used with the teachers.

Student Interviews

To add depth to quantitative data collected from students (Phases 1 and 4 of evaluation model), in-depth, semi-structured interviews were used. These interviews were conducted with students in one of the classes of each of the six case study teachers after each of the classroom observations. The interview questions were based largely on the items and scales in the WIHIC and attitude scale. Interviews were conducted in the students' mother tongue by myself. Each interview was audio recorded and later translated into English, paraphrased and transcribed for analysis.

3.6 Data Analysis

This section describes the analysis for both the quantitative and qualitative data collected during the study. The quantitative data analysis is discussed in Section 3.6.1 and the analysis of the qualitative data is described in Section 3.6.2.

3.6.1 Analysis of Quantitative Data

Once the administration of the surveys was complete, I checked and matched each student's responses to ensure that only the data for students present for both the pre-test and post-test were included. All incomplete or unusable questionnaires were discarded and the responses for the remaining questionnaires were entered into Microsoft Excel spread sheet. The responses of Almost Never, Seldom, Sometimes, Often and Almost Always were entered as 1, 2, 3, 4 and 5 respectively. Other information, including the teacher's name, classroom and grade level was

also entered at this time. To check for errors, the researcher selected a random sample of the surveys and compared them with the data entered into the database.

A total of 2,417 usable questionnaires were entered (the sample for which described in Section 3.4). This section describes the procedures used to prepare and analyse the quantitative data (described in Section 3.6.1) and the qualitative data (described in Section 3.6.2).

Quantitative data analyses were completed using SPSS version 17 statistical package. This section describes the statistical analyses conducted to address each research objective.

3.6.1.1 Validity and Reliability of the WIHIC and Attitude Scale

To examine the reliability and validity of the WIHIC when used with high school students in Indonesia (Research Objective 2), factor analysis, alpha Cronbach reliability, discriminant validity and ability to differentiate between the perceptions of students in different classrooms were examined.

As a first step, factor analysis was used to determine whether the 56 items of Indonesian version of the WIHIC measured seven independent dimensions of the learning environment. Principal axis factoring with oblique rotation was considered to be the most appropriate type of factor analysis because it can be assumed that the factors within the learning environment are related (Costello & Osbourne, 2005; Coakes & Steed, 2007). Analysis was conducted separately for the two administrations namely, one before the professional development began (pre-test) and the one near the end of the school year (post-test).

To determine whether the items in each of the seven scales and one attitude scale assessed a similar construct, internal consistency reliability for the WIHIC was calculated using the Cronbach alpha coefficient.

Discriminant validity is achieved when the correlations between a particular item and other items in the same construct are higher than its correlation with items from different constructs. To examine the discriminant validity of the scales of the Indonesian WIHIC, the component correlation matrix, obtained from oblique rotation, was used to provide an indication of whether the correlation values met the criteria of discriminant validity.

A one-way analysis of variance (ANOVA), with class membership as the independent variable, was used to examine whether the scales of the WIHIC could distinguish between classes.

For the Enjoyment scale of the TOSRA, the Cronbach alpha reliability was calculated to provide an estimate of the internal consistency. The Cronbach alpha coefficient was calculated at two units of analysis, individual and class mean.

3.6.1.2 Effectiveness of the MGMP Empowerment Programme

To evaluate the effectiveness of the MGMP Empowerment professional development programme in terms of changes in students' scores on the learning environment and attitude scales (before and after the teachers participated in the programme) (Research Objective 3), one-way multivariate analyses of variance (MANOVA) and effect sizes were used. The seven WIHIC scales and one attitude scale were used as the dependent variables and the testing occasion (pre-test and post-test) were used as the repeated measures (William, 1998). In addition, the effect size for each scale was calculated, to determine whether these differences might be educationally significant (as recommended by Thompson, 2002). The effect size was calculated by dividing the difference between the two means by the pooled standard deviation, thus expressing the difference in the means of the two samples in standard deviation units.

3.6.1.3 Differential Effectiveness of the MGMP Empowerment Programme for Teachers in Urban and Rural Schools

To explore differences and interactions between location (rural and urban students' perceptions of the learning environment) and testing occasion (before and after the teacher professional development) for each learning environment and students attitude scale (Research Objective 4), two-way MANOVA were used. The seven learning environment scales of the modified WIHIC, and one attitude scale, were used as the set of dependent variables. The two independent variables were the location (urban and rural) and testing occasion (pre-test and post-test). In particular, the interaction of the testing occasion and location provided information about the differential effectiveness of the MGMP Empowerment programme for teachers in urban and rural areas.

3.6.2 Qualitative Data Analysis

The inclusion of qualitative data in the mixed-methods model used in this study was two-fold. First, qualitative data were used to provide depth and understanding to quantitative results. Second, the qualitative data enabled me to more fully understand the MGMP empowerment programme and to provide a more in-depth evaluation of its implementation at the classroom level. The qualitative data were gathered using teacher reflective journals, interviews and classroom observations.

The qualitative data analysis involved coding the data, dividing the text into small units (phrases, sentences, or paragraphs), assigning a label to each unit, and then grouping the codes into themes. The core feature of my qualitative data analysis was the coding process. Coding was used to group evidence and to label ideas so that they reflected increasingly broader perspectives. My evidence was grouped into codes and codes were then grouped into broader themes. Themes then were grouped into larger dimensions or perspectives (Yin, 2011). The coding labels came from the exact words of the teachers, phrases composed by the researcher and concepts used in the social or human sciences (Creswell & Plano, 2010).

The data analysis is discussed under the following subheadings: the internal validity measurement of qualitative data collected (Section 3.6.2.1), analysis of qualitative information (Section 3.6.2.2) and triangulation (Section 3.6.2.3).

3.6.2.1 Internal Validity Measurement

To assess the degree to which a study accurately reflects or assess the specific concept or phenomenon that I was attempting to measure, it is necessary to check the validity. There are six basic strategies that an investigator can use to ensure internal validity for qualitative research (Merriam, 2002), these are triangulation, member checking, long-term observations, peer examination, participatory modes of research, and clarifying the researcher's assumptions. In the present study, member checking was used to ensure the internal validity of teacher reflective journal, interviews and classroom observations.

Teacher Reflective Journal and Interviews

To ensure the internal validity of the findings related to teacher reflective journal and interviews, I used member checking. I first transcribed the reflective journals and interviews and then gave the transcripts to the respondents to confirm and verify. During this process, I was able to clarify the terms that were used in the transcripts. In addition, my interpretations of respondents' comments or views were also returned to the respondents to further confirm and enhance the credibility of the data, as suggested by Kvale (1996) and Rubin (2005).

Classroom Observations

To ensure the internal validity of classroom observations, the following procedure was applied. Before classroom observations were conducted, the six case study teachers were informed that the researcher would act as a non-participant observer and would not make personal value judgements about the quality of their teaching practices. This effort was taken to ensure classroom transactions occurred in a more natural manner. The observations were video recorded for further analysis. Member checking was used to ensure the validity of the data collected during observations. After each observation, interviews with the teachers and students

regarding the events that were observed during the classroom observations were used to help to clarify the researcher's judgement. In this way the participants helped to ensure the validity of that data collected from classroom observations. After each classroom observation, I also discussed the notations that I had made with respect to my views of the observations, with the teacher who was observed as an additional accuracy check (Merriam, 2002).

3.6.2.2 Analysis of Qualitative Information

To analyse the data gathered using teacher reflective journals and interviews, I used framework (thematic) analysis (Ritchie & Lewis, 2003). This form of analysis allows for the inclusion of *a priori* as well as emergent concepts. By using NVivo 8 software, the reflective journals of the six case study teachers were coded as follows. First, I entered the transcript documents directly into NVivo package. Second, comments were grouped into pre-existing categories (nodes) that were adopted from Guskey's (2000) framework. Items that had been coded were reread to look for consistencies and to clarify themes. I reviewed the data and discussed the patterns of themes that had been found with my supervisor.

The procedures for analysing the interview data were orderly; first, the interview recordings were transcribed into English. Second, information identified from the transcript was categorised and structured to create a coherent explanation or description of aspects that contributed to the teachers' and students' perceptions based on the frequency. Finally, the transcripts were coded and analysed in NVivo 8 software package, by identifying the key terms (provided by Guskey's framework) that were used during the interview.

Within the context of the study, the video data of classroom observations were analysed by implementing a cyclical analytical process (Jacobs, Kawanaka & Stigler, 1999) that included watching, coding and analysing, with the goal of transforming the video images into objectives and verifiable information. During this process, I watched the tapes and generated hypotheses related to the research objectives. Once hypotheses were generated, I developed a coding system to test the ideas.

The coding system required me to watch repeated viewings of particular tapes. From this cycle, I developed codes of particular segments of video. In the next step, I applied the coding system to corroborate or disconfirm my initial hypotheses.

3.6.2.3 Triangulation

The data collected from the classroom observations, interviews and teacher reflective journals were triangulated in the construction of case studies as means of telling a richer story of the classroom. Triangulation is a one-phase design in which researchers implement the quantitative and qualitative methods during the same timeframe and with equal weight (Creswell & Plano Clark, 2010). The most well-known approach to mixed-methods design is the triangulation design, which is often referred to as a convergent parallel design (Jick, 1983). Jick (1983) stated that triangulation is used to achieve coverage of various aspects of a phenomenon of interest and added that, using qualitative and quantitative methods together may also illuminate elements of the context of a phenomenon, thereby achieving a more complete understanding.

Within the context of the study, I developed case studies by merging the two data sets (quantitative and qualitative). Development of the case studies incorporated not only my observations of the impact of the MGMP teacher professional programme on the classroom level but also the conversations between the teachers and myself and from the classroom learning environment data based on the Indonesian version of the WIHIC and attitude scale.

3.7 Ethical Considerations

Before conducting my research, I requested the approval from the bureaucracy in the Department of Education. Initially, the Head of Department of Education in Central Java province was contacted and permission was requested to conduct the research in the district schools within his province. The Head of Department of Education of Central Java province granted permission, after which, the Heads of Department of Education in the six districts within Central Java province were

contacted and permission was requested to conduct the research in the schools within their regencies. All the six of the Heads of Department of Education granted permission.

Finally, the principals of 32 schools from which the participating teachers were teaching were contacted for permission. In line with the ethics guidelines for research at Curtin University (Curtin University, 2009), an information sheet and research consent form were distributed to the participating principals, teachers, students and their parents. The information sheet outlined: the goals of the research; the role of the teachers; consent to participate; and the assurances of anonymity and confidentiality.

Consent to participate suggested that the involvement in the research was entirely voluntary. The principals, teachers, students and their parents had right to withdraw at any stage of the research without it affecting the teachers' right or responsibility.

Anonymity was assured and personal information about the teachers was not identified in the report. Confidentiality assured that the information the teachers provided was kept with the utmost confidentiality. All names were kept separate from their personal details, and only the researcher and her supervisor would have access to this. All of the data collected did not have any names or any other identifying information on them, in adherence to university policy.

3.8 Chapter Summary

This chapter has presented the research methods used in the present study. As a first step, to answer research question one, a model that could be used to evaluate the teacher professional development programme in the Indonesian context, was developed. The new model included the salient aspects of three existing models developed by Fishman, Mark, Best & Tal (2003), Guskey (2000) and Mathison (1992) and incorporated the assessment of the students' learning experiences to provide a

holistic evaluation that suited to the Indonesian context. The model involves five phases which were used to guide the collection and analysis of the data.

The selection of my sample employed both multiple cluster sampling and purposive sampling. The sample for the collection of quantitative data involved 2,417 students in 66 classes who were present for the both the pre-test (collected at the start of the programme) and the post-test (collected one year later at the end of the programme). In addition, qualitative data were collected from 33 teachers and six teachers who were to form the basis for case studies.

To evaluate the professional development programme, the study used a mixed-methods approach (guided by the evaluation model) that involved the collection and combination of quantitative and qualitative data. The quantitative component involved a pre–post design in which two surveys were administered to determine whether there were changes in students' perceptions of their learning experiences and their enjoyment of English classes.

After careful consideration, I selected and modified a widely-applicable perceptual measure, the What Is Happening In this Class? (WIHIC) questionnaire and an attitude scale based on the Test of Science-Related Attitude (TOSRA). To ensure the suitability of the instruments to the Indonesian context, they were modified. First, one scale of the WIHIC, Investigation, was omitted as it was considered irrelevant to English language classes. Second, the Finding Reference scale was developed as this formed an important aspect of the professional development programme that was not included in the WIHIC. Third, individual items of the WIHIC were scrutinised for suitability and minor changes were made. To ensure that the attitude scale was suitable, several modifications were made to the items by changing the wording of all items from 'Science Lessons' into 'English Lessons'. Also, to avoid confusion, the response format was changed to match that used in the WIHIC. This also required some changes to the wording of individual items.

To translate the two instruments into Bahasa Indonesia, the process of back translation was used. The process involved the English version being translated into Bahasa Indonesia by myself and then translated back into English by an independent party who was fluent in both languages but not familiar with the survey. This allowed the comparison of the two English versions for equivalence in meaning.

Once the instruments were modified and translated, a pilot test involving 89 students (not involved in this study) was conducted. The pilot study served to ensure that individual items of the surveys were clear and unambiguous and that the students comprehended them in the way in which the researcher intended. The pilot study was also used to ascertain how long it took the students to complete the surveys and to ensure that the instructions were clear.

The qualitative component employed a case study approach that involved the gathering of information from a variety of sources including teacher reflective journals, in-depth interviews with teachers and students and classroom observations. The qualitative information helped to explain, clarify and contextualise the quantitative overview, whilst also providing causal information for the quantitative results.

As a first step, the reliability and validity of the WIHIC and attitude scales (Research Objective 2) were examined to ensure that the data could be used with confidence. To examine the reliability and validity of the instruments a variety of statistical analyses, including, factor analysis, alpha Cronbach reliability, discriminant validity and ability to differentiate between the perceptions of students in different classrooms, were used.

Research Objective 3 involved evaluating the effectiveness of the MGMP Empowerment programme. The evaluation model (developed at Research Objective 1) was used to guide the collection and analysis of data during the evaluation. To examine teachers' views of the relevance and utility of the teacher professional

development (Phase 2 of the evaluation model), teachers' reflective journals and in-depth interviews with teachers were used. To examine the extent to which teachers made changes to their classroom practices, in ways that were consistent with the pedagogical ideas imparted during the professional development programme, observations were conducted in each of the six case study teachers' classrooms before the commencement of the programme (Phase 1 of the evaluation model) and again towards the end of the programme (Phase 3 of the evaluation model). To evaluate the programme in terms of changes in students' perceptions of their learning experience and attitudes, one-way MANOVA and effect sizes were used (Phase 4 of the evaluation model). Because observations carried out at Phase 3 of the evaluation model indicated that there could be differences in the effectiveness of the programme for teachers in urban and rural schools, two-way MANOVA and effect size were used to explore the differential effectiveness of the programme. Finally, in-depth interviews with the six case study teachers were used to evaluate whether there were contextual factors that influenced the implementation of the MGMP Empowerment programme (Phase 5 of the evaluation model).

The data gathered using teacher reflective journals and interviews were analysed using framework (thematic) analysis in which coding was the core feature of the analysis. The video recordings of classroom observations were analysed by implementing a cyclical analytical process. The triangulation of the various data sets, in the construction of case studies was used to tell a richer story of the teachers' attempts to translate the programme into practice.

To ensure the internal validity of the findings, the data gathered using interviews, member checking was used. All interviews were transcribed and the transcripts were given to the respondents to confirm and verify. Similarly, observation checklists and field notes were discussed with all teachers to verify that the recordings were correct and to ensure the internal validity of data.

The results of the data analyses are discussed in Chapters 4 and 5. Chapter 4 focuses on the validity and reliability of the instruments used my study while Chapter 5 focused on the evaluation of the MGMP Empowerment programme.

Chapter 4

ANALYSIS AND RESULTS:
Validity and Reliability of the Surveys

4.1 Introduction

The overarching purpose of the present study was to evaluate the effectiveness of the MGMP Empowerment programme for lower secondary English teachers. An important component of this evaluation was the use of two instruments: one to assess students' perceptions of their learning environment (based on the What Is Happening In this Class? (WIHIC) questionnaire) and one to assess students' attitudes towards their English classes (based on the Test Of Science-Related Attitudes (TOSRA)). These two instruments were administered once at Phase 1 (pre-test) and again, one year later, at Phase 4 (post-test) of the evaluation model. A description of both of these instruments, their modification to suit the Indonesian context and their translation were detailed in Chapter 3.

Before using the instruments as part of my evaluation, it was important to establish their reliability and validity to allow increased confidence in drawing conclusions based on the data collected using the instruments. Thus this chapter is devoted to answering the second research objective:

To modify, translate and validate Indonesian versions of two surveys (one to assess students' perceptions of the learning environment and one to assess students' attitudes towards their English classes) for use at the lower secondary school level in Indonesia.

This chapter reports the reliability and validity of the modified WIHIC (discussed in Section 4.2); and the reliability and validity of the attitude scale (discussed in Section 4.3).

4.2 Validity and Reliability Analyses of the WIHIC

In this study, the students' perceptions of their classroom learning environment were assessed using the What Is Happening In this Class? (WIHIC, Fraser, Fisher & McRobbie, 1996). The data collected from the sample of 2,417 students in 66 classes in 32 lower secondary schools (described in Chapter 3), were analysed to examine the validity of the WIHIC in terms of: the factor structure (Section 4.2.1); internal consistency reliability of the scales (Section 4.2.2); the discriminant validity (Section 4.2.3); and each scale's ability to differentiate between classes (Section 4.2.4).

All analyses were conducted separately for the data collected from the two administrations of the questionnaire, the pre-test (conducted before the teacher professional development programme began) and the post-test (conducted at the end of the school year). The purpose of the pre and post-test administrations was to permit evaluation of the programme in terms of changes in students' perceptions of their learning environment and their attitudes towards their English classes.

4.2.1 Factor Structure

The data collected for the pre-test and post-test involved 2,417 students all of whom were present for both. These data were analysed to examine the internal structure of the Indonesian version of WIHIC. The establishment of meaningful groups of items, identified by examining the factor loadings, is important in ascertaining validity with factor analysis (Tabachnick & Fidell, 2001). Principal axis factor analysis with oblique rotation was used to generate orthogonal factors for each of the two data sets (pre-test and post-test). I selected oblique rotation because the dimensions of the learning environment can be considered to overlap

(Costello & Osborne, 2005; Floyd & Widaman, 1995). In this study, I used the conventionally-accepted cut-off of 0.30 for factor loadings (Stevens, 2002).

Table 4.1 reports the results separately for the two administrations of the Indonesian version of the WIHIC (pre-test and post-test). The results indicated that, of the 56 items, eight items were problematic: Items 6 and 8 from the Student Cohesiveness scale, Item 16 from the Teacher Support scale, Items 17, 19 and 23 from the Involvement scale, Item 28 from the Finding Reference scale and Item 43 from the Cooperation scale. These items were all omitted to strengthen the factor structure and were not included in any further analyses.

The factor analysis resulted in the acceptance of a revised version of the instrument comprising of 48 items in the seven *a priori* scales of Student Cohesiveness, Teacher Support, Involvement, Finding Reference, Task Orientation, Cooperation and Equity. All of the items in the revised instrument had a factor loading of at least 0.30 on their *a priori* scale with the exceptions of two of items, both on the pre-test: Item 7 from the Student Cohesiveness scale (which did not load on its own scale but loaded on the Cooperation scale) and Item 49 from the Equity scale (which did not load on its own scale but loaded on the Student Cohesiveness scale).

The percentage of variance extracted with each factor (see Table 4.1) for the pre-test varied from 3.61% to 26.41% for different scales, with the total variance accounted for being 56.39%. For the post-test, the percentage of variance ranged from 3.36% to 29.27 % for different scales, with a total variance accounted for being 59.67%. The values of the eigenvalue varied from 1.3 to 12.71 for different scales for the pre-test and from 1.18 to 14.63 for the post-test.

The results of the factor analysis for the 48-item version of the WIHIC are consistent with the results of past studies that were conducted in Indonesia (Aldridge, Fraser & Adolphe, 2010; Margianti, Aldridge & Fraser, 2004; Wahyudi & Treagust, 2004), Singapore (Chionh & Fraser, 2000), the United States (Allen & Fraser, 2007; Ogbeuhi & Fraser, 2007; Wolf & Fraser, 2008), the United Arab Emirates (Afari, Aldridge,

Fraser & Khine, in press; MacLeod & Fraser, 2010), Australia and Taiwan (Aldridge & Fraser, 2000; Aldridge, Fraser & Huang, 1999).

4.2.2 Internal Consistency Reliability

To examine the extent to which the items in each scale assess a common construct, internal consistency reliability was used. Cronbach's (1970) alpha coefficient, using both the individual and the class mean as the units of analysis, was calculated separately for both the pre-test and post-test administrations to provide an estimate of the internal consistent reliability for each scale. Table 4.2 reports the results for the seven scales of the Indonesian version of the WIHIC, for both the pre-test and post-test.

The alpha coefficients for the seven scales of the Indonesian version of WIHIC, for the individual student as the unit of analysis, ranged from 0.69 to 0.89 for the pre-test and from 0.73 to 0.91 for the post-test. With the class mean as the unit of analysis, the alpha reliability coefficients were generally higher, ranging from 0.91 to 0.98 for the pre-test and 0.92 to 0.98 for the post-test. According to McMillan and Schumacher (2001), in the case of exploratory research, a reliability of 0.50 is acceptable and a Cronbach alpha reliability coefficient between 0.70 and 0.90 is considered to be high. As the Cronbach alpha coefficient for each of the seven scales for both units of analysis, for both the pre-test and post-test administrations (with the exception of Student Cohesiveness for the pre-test) were all above 0.70, they can be considered to be high according to McMillan and Schumacher's (2001) criteria. These results are consistent with previous studies and, in particular, it replicates the results of a previous study involving an Indonesian version of WIHIC in secondary school students (Aldridge et al., 2010).

Table 4.1 Factor Loadings for WIHIC items

Item No	Factor Loading													
	Student Cohesiveness		Teacher Support		Involvement		Finding Reference		Task Orientation		Cooperation		Equity	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	0.58	0.70												
2	0.51	0.52												
3	0.44	0.42												
4	0.62	0.65												
5	0.32	0.33												
7	–	0.31									0.38			
9			0.60	0.64										
10			0.69	0.66										
11			0.52	0.58										
12			0.62	0.57										
13			0.49	0.56										
14			0.56	0.60										
15			0.39	0.41										
18					0.45	0.44								
20					0.40	0.46								
21					0.37	0.38								
22					0.49	0.55								
24					0.42	0.45								
25							0.44	0.60						
26							0.48	0.60						
27							0.63	0.71						
29							0.63	0.74						
30							0.65	0.71						
31							0.65	0.67						
32							0.64	0.60						
33									0.56	0.49				
34									0.62	0.58				
35									0.57	0.54				
36									0.64	0.66				
37									0.62	0.59				
38									0.65	0.69				
39									0.63	0.67				
40									0.57	0.57				
41											0.61	0.62		
42											0.44	0.44		
44											0.39	0.39		
45											0.63	0.62		
46											0.81	0.85		
47											0.73	0.77		
48											0.46	0.42		
49	0.63												–	0.69
50													0.68	0.70
51													0.74	0.72
52													0.74	0.77
53													0.72	0.75
54													0.72	0.72
55													0.64	0.65
56													0.69	0.70
% Variance	4.47	5.01	3.67	3.77	3.61	3.36	6.94	29.27	5.61	5.72	5.41	5.43	26.41	7.11
Eigen-value	1.8	2.01	1.34	1.39	1.30	1.18	5.94	14.63	2.31	2.36	2.21	2.21	12.71	3.05

Factor loadings less than 0.30 have been omitted.

N= 2417 students in 66 classes for each of the pre-test and post-test

Table 4.2 Internal Consistency Reliability (Cronbach Alpha Coefficient) for Two Units of Analysis for the Modified WIHIC and Attitude Scale

Scale	Unit of Analysis	No of Items	Alpha Reliability	
			Pre-test	Post-test
Learning Environment				
Student Cohesiveness	Individual	6	0.69	0.73
	Class Mean		0.91	0.92
Teacher Support	Individual	7	0.82	0.85
	Class Mean		0.94	0.96
Involvement	Individual	5	0.75	0.79
	Class Mean		0.93	0.96
Finding Reference	Individual	7	0.82	0.88
	Class Mean		0.95	0.98
Task Orientation	Individual	8	0.85	0.86
	Class Mean		0.97	0.97
Cooperation	Individual	7	0.82	0.84
	Class Mean		0.94	0.95
Equity	Individual	8	0.89	0.91
	Class Mean		0.98	0.98
Attitude				
Enjoyment of English Classes	Individual	10	0.90	0.91
	Class Mean		0.97	0.98

N= 2417 students in 66 classes at 32 schools

4.2.3 Discriminant Validity

Brown (2006) and Field (2009) explained that oblique rotation in exploratory factor analysis provides realistic representation of how factors are interrelated. According to Field (2009), based on theoretical grounds, there should be a moderately strong relationship between factors. However, if the correlation between the factors of an instrument is above 0.80 then it may imply overlap of concepts and point towards poor discriminant validity (Brown, 2006). The results of what for the components correlation matrix, generated using Oblique rotation, for both the pre-test and post-test suggest that the Indonesian version of the modified WIHIC scales measure distinct constructs for both administrations (see Table 4.3). The highest correlation between the scales of the WIHIC was 0.46 and this value met the requirements of

discriminant validity, according to Brown (2006). The results were similar to past studies used the WIHIC which found that the scales were able to measure distinct constructs (Biggs, 2008; Velayutham, Aldridge & Fraser, in press).

Table 4.3 Component Correlation Matrix for WIHIC for Pre-test (Correlations above the Diagonal) and Post-test (Correlations below the Diagonal)

Scale	Correlations						
	Student Cohesiveness	Teacher Support	Involvement	Finding Reference	Task Orientation	Cooperation	Equity
Student Cohesiveness	–	0.35	0.41	0.36	0.20	0.36	0.30
Teacher Support	0.35	–	0.46	0.29	0.20	0.18	0.43
Involvement	0.41	0.46	–	0.30	0.31	0.24	0.28
Finding Reference	0.36	0.29	0.30	–	0.28	0.29	0.27
Task Orientation	0.20	0.20	0.31	0.28	–	0.18	0.25
Cooperation	0.36	0.18	0.24	0.29	0.18	–	0.33
Equity	0.30	0.43	0.28	0.27	0.25	0.33	–

4.2.4 Ability to Differentiate between Classrooms

An analysis of variance (ANOVA), with class membership as the independent variable, was used to examine the extent to which students within the same class perceived the classroom environment relatively similarly to each other and to which the mean class perceptions varied from class to class. The ANOVA results, reported in Table 4.4, indicate that each WIHIC scale differentiated significantly (p<0.01) between classrooms for both the pre-test and the post-test data sets. The eta^2 statistics, which represents the proportion of the variance in environment scores accounted for by class membership, ranged from 0.19 to 0.25 for the pre-test and from 0.21 to 0.30 for the post-test. Overall, the ANOVA results provide further evidence that the learning environment scales of the Indonesian version of the modified WIHIC are able to differentiate between classrooms. The results of this study replicate numerous past studies, which have consistently found that scales of the WIHIC were able to differentiate significantly (p<0.01) between students'

perceptions in different classrooms (Aldridge at al., 2010; Aldridge & Fraser, 2000; Allen & Fraser, 2007; Khoo & Fraser, 2008).

Table 4.4 Ability of the WIHIC to Differentiate Between Classrooms (ANOVA Results) for the WIHIC

Scale	ANOVA (eta^2)	
	Pre-test	Post-test
Student Cohesiveness	0.25**	0.30**
Teacher Support	0.23**	0.30**
Involvement	0.19**	0.28**
Finding Reference	0.19**	0.25**
Task Orientation	0.20**	0.21**
Cooperation	0.20**	0.23**
Equity	0.25**	0.27**

N= 2417 students in 66 classes in 32 schools
** $p< 0.01$ * $p< 0.05$
The eta^2 statistic (which is the ratio of 'between' to 'total' sums of squares) represents the proportion of variance explained by class membership.

The factor structure, scale internal consistency reliability, discriminant validity and ability to differentiate between classrooms, support the reliability and validity of the Indonesian version of WIHIC when used in high school English classes in Indonesia. Therefore, it can be concluded that the data collected using the modified WIHIC can be used with confidence to examine the remaining research objectives.

4.3 Validation of Attitude Scale

To assess students' attitudes towards English classes, I developed an attitude scale, modelled on the Enjoyment of Science Lessons scale from the Test of Science-Related Attitude (TOSRA, Fraser, 1981) as described in Chapter 3. Past studies have successfully modified the TOSRA for use in a range of countries, including in the United States (Spinner & Fraser, 2005), Indonesia (Aldridge, Fraser & Adolphe, 2010; Margianti, Aldridge & Fraser, 2004) and Taiwan and Australia (Aldridge, Fraser & Huang, 1999).

The internal consistency reliability was calculated for the attitude scale used in my study using Cronbach's alpha coefficient. The results reported in Table 4.2 show that the Cronbach alpha reliability for the attitude scale, for both the pre-test and the post-test was high. With the student as the unit of analysis, the alpha coefficient was 0.90 for the pre-test and 0.91 for the post-test. Using the class unit of analysis, the alpha coefficient was 0.97 for the pre-test and 0.98 for the post-test.

Taken together with the reliability results for the Indonesian version of WIHIC, these results suggest that both attitude and learning environment scales were reliable when used high school students in English classrooms in Indonesia. Therefore they could be used with confidence to examine the other research objectives in my study.

4.4 Chapter Summary

Part of my study involved evaluating the effectiveness of the MGMP Empowerment programme in terms of changes in students' perceptions of their classroom learning environment and their attitudes towards English. It was important, therefore, to ensure that the instruments were valid and reliable in the Indonesian context because both instruments were originally developed in Western countries and modified and translated into Indonesian for the purpose of this study.

My research involved both a pre-test (conducted at Phase 1 of the evaluation) and a post-test (conducted at Phase 4 of the evaluation). Both phases involved administration of both of the questionnaires, one to assess students' perceptions of their classroom learning environment (WIHIC) and another to assess students' attitudes modelled on a scale from the TOSRA.

The data collected from 2,417 students (for both the pre-test and post-test) in 66 classes were analysed to examine the validity and reliability of the WIHIC scales (namely, Student Cohesiveness, Teacher Support, Involvement, Finding Reference, Task Orientation, Cooperation and Equity) and the attitude scale (namely,

Enjoyment for English Lesson). For the WIHIC, factor structure, internal consistency reliability, discriminant validity, and ability of the learning environment scales to differentiate between classrooms for both the pre-test and post-test administrations were used as indices of validity. For the attitude scale, the internal consistency reliability was used to determine its reliability.

For the modified WIHIC, principal axis factoring with oblique rotation confirmed a refined 48 items version. For this 48-item version of the WIHIC, all items had a factor loading of at least 0.30 on their *a priori* scale and no other scale, with the exception of two items in the pre-test (one the Student Cohesiveness scale and one in the Equity scale).

The Cronbach alpha reliability coefficients ranged from 0.69 to 0.90 for both pre-test and post-test administrations of the WIHIC, for both the individual and the class mean as the unit of analysis. The results of one-way ANOVA were statistically significant for each scale, indicating that each of the WIHIC scales was able to differentiate between the perceptions of the students in the different classrooms. Values in the component correlation matrix obtained from the oblique rotation were above 0.18 and below 0.46 for all scales, thereby supporting the discriminant validity of the WIHIC scales.

To determine the reliability of the attitude scale, the Cronbach alpha reliability was used as a convenient index. The alpha reliability coefficient for the attitude scale, with the student as the unit of analysis, the alpha coefficient was 0.90 for the pre-test and 0.91 for the post-test. Using the class unit of analysis, the alpha coefficient was 0.97 for the pre-test and 0.98 for the post-test.

Overall, the results supported the validity and reliability of the seven WIHIC scales and the attitude scale when used with lower secondary school students in Central Java, Indonesia. These results were, by and large, acceptable and were comparable to those reported in previous studies (Aldridge et al., 2002; Fraser, Aldridge & Adolphe, 2010; Margianti & Fraser, 2010; Wahyudi & Treagust, 2004).

Based on the results reported in this chapter, the 48-items Indonesian version of WIHIC and the attitude scale, modelled on TOSRA, were considered to be valid and reliable for assessing the learning environment and students' attitudes in Indonesian secondary lower schools. It was justifiable, therefore, to use the data for further analysis.

Chapter 5 discusses the results of my evaluation of the MGMP Empowerment programme.

Chapter 5

ANALYSIS AND RESULTS

Effectiveness of the MGMP Empowerment Programme

5.1 Introduction

In the previous two chapters I reported the research methods that I used to evaluate the MGMP Empowerment programme (Chapter 3) and the reliability and validity of the student surveys that were used as part of the evaluation process (Chapter 4). This chapter reports the analysis and results for the data collected at various phases of the evaluation model. The evaluation model, described in Section 3.2, involved five distinct phases. The first phase involved the collection of qualitative and quantitative baseline data. The second phase involved examining the teachers' views of the professional development experience and their views about whether this experience had increased their knowledge and skills with respect to teaching English. The third phase involved examining changes in classroom practice as teachers used of the knowledge and skills that were the focus of the professional development programme. The fourth phase involved evaluating the impact of the professional development programme in terms of changes in students' perceptions of their learning experience and attitudes. Finally, the fifth phase involved examining of contextual factors that may have promoted or impeded the translation of the professional development programme into practice.

The results are organised to coincide with the framework of the model. Given that the first phase of evaluation involved the collection of baseline data, these are reported at phases three and four. The following headings are used to report the evaluation:

- Teachers' Views of the Relevance and Utility of the Teacher Professional Development Programme (Section 5.2);

- Teachers' Use of Knowledge and Skills (Section 5.3);
- Changes in Students' Perceptions of the Learning Experience and Attitudes (Section 5.4);
- Contextual Factors that Promote or Impede the Programme Implementation (Section 5.5); and
- Chapter Summary (Section 5.6).

5.2 Teachers' Views of the Relevance and Utility of the Teacher Professional Development Programme (Phase 2)

To help to understand teachers' views of the usefulness and relevance of the MGMP Empowerment programme, reflective journals and in-depth interviews were used. First, reflective journals written by the teachers (*N*=138) who attended the professional development programme, were used to elicit the general themes of the analysis, after which, the reflective journals of six case study teachers were used for further analysis. Second, in-depth semi-structured interviews were conducted with each of the six case study teachers, to help to understand their views about the usefulness, relevance and impact of the programme. (See Chapter 3 for details related to the sample and data collection methods used.) The results of this phase of the evaluation are organised into two subheadings: teachers' views of the professional development experience (discussed in Section 5.2.1); and teachers' views about whether the professional development expanded their knowledge and skills (discussed in Section 5.2.2).

5.2.1 Teachers' Views of the Professional Development Experience

This phase of the evaluation examined teachers' views of the quality of the professional development experience (Fishman et al., 2003; Guskey, 2000; Mathison, 1992; Kirkpatrick, 2005), thus, providing information about the teachers' reactions to the professional development experience. Information for this phase

was gathered using in-depth interviews with teachers and teacher reflective journals.

This phase of the evaluation focused on teachers' attitudes towards a range of aspects related to the delivery of the MGMP Empowerment programme. The themes distilled during the analysis were used to describe how the teachers viewed their experiences, what the teachers regarded as beneficial and, importantly, how the programme could be improved in the future. The results of analyses are reported in terms of the context (i.e. facilities, accommodation and refreshments), in Section 5.2.1.1; and the delivery of the programme (i.e. the quality of the programme leaders, time available and the activities used), in Section 5.2.1.2.

5.2.1.1 Context

The environment in which the professional development experience takes place is considered to be important. According to Guskey (2000), attending to basic human issues in a professional development programme, such as making sure that teachers are comfortable and that the environment is conducive to learning, will do much to enhance teachers' regard for the experience.

Analysis of the data indicated that all of the teachers, without exception, were positive about the facilities that were provided. To this end, one of the teachers commented, "I think that, in general, the teacher professional development programme was good, especially with respect to the facilities provided in the programme; such as the accommodation, training rooms, media and refreshments". Similarly, another teacher stated, "the facilities provided for the programme were very good, especially the accommodation which was newly refurbished".

5.2.1.2 Delivery

When assessing the delivery of the MGMP Empowerment programme, the interviews with the teachers (*n*=33) indicated that there were two aspects which may have hampered the success of the professional development. First, the calibre of two of the four trainers and, second, the tight time schedule required to

complete the programme. All of the teachers whom were interviewed felt that two of the four trainers were not adequately prepared. These teachers agreed that the presentations given by these two trainers failed to provide them with clear teaching suggestions and strategies, or to explain the teaching materials adequately. According to these teachers, the workshops provided by these two trainers were little more than one-way lectures, leaving the teachers with the impression that they had inadequate knowledge of the curriculum and lacked the appropriate teaching skills. To this end, one of the case study teachers commented, “The presentations [of these two trainers] were not satisfactory, the workshops were one-way lectures and the tutors did not explain the materials clearly”. Another teacher commented, “I am sure that two of the trainers did not have enough knowledge about the new English curriculum and its implementation in English classes”.

All of the teachers whom were interviewed (*n*=33) also complained that the amount of time available to cover the content of the programme. The professional development programme was carried out over four days each semester, with each day starting at 8 o’clock in the morning and finishing at 10 o’clock in the evening. Teachers’ interviews and entries in their reflective journals indicated that the majority of them suffered from fatigue and that the programme felt rushed. Many complained that the tight programme left little time for discussions and that they felt overloaded with information. To this end, one teacher stated, “The schedule of the programme was not effective and the workshop was rushed. It was not effective because we were very tired and, of course, we were not productive as we were ‘overloaded with so many materials’ in such a short amount of time”.

A study conducted by Ingvarson, Meiers and Beavis (2005) indicated that time span and contact hours of professional development was important in the design of effective professional development. In addition, Garet et al. (2001) claimed that professional development was likely to be higher quality if it is both sustained over time and involved a substantial number of hours.

Based on entries in the reflective journals and interviews with teachers, the results indicated that although the teachers reacted positively to the context of the professional development, they were critical about two aspects of the delivery of the MGMP Empowerment programme. These initial reactions to their experience, that is their satisfaction with the experience, were important findings in terms of future improvement for the programme.

5.2.2 Teachers' Views about Whether the Professional Development Increased Their Knowledge and Skills

Ideally, professional development programmes should be carefully planned and conducted in ways that ensure that all those who take part consider it to be beneficial. However, professional development should do more than simply make teachers feel good, it also should be a learning experience for all who are involved. Therefore, an important component of this evaluation was to examine teachers' views about the extent to which the professional development programme had increased their knowledge and skills. To do this, I analysed teacher reflective journals and the interview data of the six case study teachers.

During in-depth interviews, all six of the case study teachers indicated that their participation in the MGMP Empowerment programme had increased their knowledge and skills. Without exceptions, they all felt that the content of the programme was relevant, practical and useful to their classroom teaching. To this end, one of the teachers commented, "The professional development was very useful and this programme will help me to improve my teaching practices which will, in turn, improve my students' learning".

During the interviews, all six of the case study teachers indicated that they were keen to improve their teaching and to develop themselves as teachers. They all agreed that improving their skills as teachers and being continuous learners were important. To this end, one teacher stated, "I think that if teachers are not motivated to develop, they will be left behind because knowledge and technology

are developing all of the time". Another teacher claimed: "I am always keen to improve the quality of my teaching. In this world, information and technology develop very quickly and teachers must make adjustments in line with these developments. It is important, therefore, that I work hard to improve my knowledge and teaching skills".

Analysis of the teachers' reflective journals indicated that all of the teachers were enthusiastic about implementing the suggestions provided during the programme. To this end, one teacher wrote: "During the programme, I received a lot of information about modern teaching strategies, content knowledge related to the English language and the application of technology in teaching and learning. I am sure that my students will become highly interested in English if I implement these suggestions in my English classes".

The case study teachers also agreed, unanimously, that the programme was a valuable opportunity for them to share ideas about teaching and learning with other teachers. The teachers found the opportunity to meet with colleagues and to share ideas and experiences was rewarding and informative. To this end, one teacher commented:

> By participating in this teacher professional development, I have had an opportunity to meet many English teachers from different cities and it is valuable because we have shared our ideas about teaching and learning and we also shared teaching resources.

These positive findings were in line with Richards and Farrell's (2005) study, which found that the teacher professional development programmes were reported by teachers to be a source of influence on their decisions, inspiring them to change their teaching and beliefs.

My analysis of data, gathered from reflective journals and interviews indicated that despite the problems related to two of the trainers, teachers felt that they had

gained knowledge and skills as a result of the professional development programme. The six case study teachers were keen to implement the ideas into their classroom practice. The results suggested that teachers found the professional development programme to be practical and useful. These results were in line with the studies conducted by Kelly (2006) and the Office for Standards in Education, London (2004) which reported that the outcomes of teacher professional development training for which there was a positive impact on teachers' subject knowledge, teaching practice, students' standards and curriculum planning skills.

The next section discusses the third phase of the evaluation which examined the extent to which the teachers changed their practices in ways that were in line with the knowledge and skills presented during the professional development programme.

5.3 Teachers' Use of Knowledge and Skills (Phase 3)

Fullan (2003) and Guskey (2002) proposed that one of the goals of teacher professional development was to change teachers' classroom practices. In evaluating the effectiveness of the teacher professional development programme, it was considered important to gather information with respect to whether or not the new skills and knowledge, taught during the programme, were being used in the classroom. My evaluation of the teachers' use of new knowledge focused on the central question: Did what the teachers learn through their professional development experience translate to changes in their professional behaviours or activities?

To provide baseline data, (Phase 1 of the evaluation) a one-hour classroom observation was held in each of the case study teachers' classrooms at the beginning of the school year. This information was used to help me to determine whether changes had taken place. During these initial classroom observations, the case study teachers all used teacher-centred methods in which one-way lectures as the predominant teaching strategy. As a major objective of the MGMP

Empowerment programme was to empower teachers to teach in ways that were more student-centred, it was anticipated that, after the one-year programme, the strategies used by the teachers would focus on student learning, rather than using traditional, didactic methods.

At Phase 3, a minimum of four hours of observations were made in each of the two classes of the case study teachers. At least one hour of these classroom observations was conducted after the MGMP Empowerment programme was finished. All of the classroom observations were video recorded and analysed by implementing a cyclical analytical process (see Section 3.6.2.2 for details) as suggested by Jacobs, Kawanaka and Stigler (1999). Interviews were held with each of the case study teachers before and after the observations. During the classroom observations, I focused on teaching behaviours and pedagogies that matched those that were suggested in the professional development programme. Given that my analysis indicated that differences existed for teachers in urban and rural schools, the following sections report the results of the classroom observations separately for changes in teaching practices in urban schools (Section 5.3.1) and changes in teaching practices in rural schools (Section 5.3.2).

5.3.1 Teaching Practices in Urban Schools

The urban schools in which classroom observations were conducted were all well-resourced. All three urban schools had a library and three language laboratories, three science laboratories, a multimedia room and a sports hall. The schools all had computers with a wireless network, that they called a 'hotspot', that was available to both teachers and students free of charge. This 'hotspot' could also be used by students and parents to access students' test and examination scores, and to check student attendance. The students of these schools came from families of middle to high socio-economic status. Most of parents worked as government officers or were successful entrepreneurs. In addition to the funding from the government, the parents paid school fees and regularly made donations to schools. There were 18 to 27 classrooms in each of the schools with25 to 35 students in each classroom.

The classrooms also were well resourced with a computer on the teacher's desk, a digital projector (hanging from the ceiling) and a screen and a whiteboard on the front wall of the classroom. In all cases, the classrooms had air conditioning and most of the students had their own laptops.

The 12 hours of classroom observations carried out in the classes in urban schools, indicated that two of the three case study teachers had made changes to their teaching practice that were in accordance with the suggestions provided in the MGMP Empowerment programme. These two case study teachers used teaching methods that were more aligned to a student-centred approach that they had not used before they attended the programme. The teachers also made use of electronic learning media (including using the computer and the internet). During my observations in the classes of these two case study teachers, the students were provided with opportunities to work in groups and, when they did, they worked well with one another. I observed the students discussing tasks with their peers in their groups and, in all cases, they worked hard to complete the group task.

Although these teachers showed evidence of more students-centred practices, it was interesting to note that the quality of these practices varied. For example, one of the observations involved the teacher starting the lesson with a revision of the previous lesson. He used a power point slide with questions related to the previous lesson, the answers to which were chorused by whole class in unison.

His instructions for the day's lesson involved a second slide that included the objectives that the students were required to achieve. A third slide involved a copy of the worksheet that the students were to complete. The students were then instructed to work in groups of four to complete the worksheet and were told that a person from each group would present their findings to the class at the end of the lesson.

The movement of the students, as they sought friends with whom they would work, was orderly and they quickly settled down to the task at hand. The students,

without exception, worked industriously to complete the worksheet. As they worked, the teacher sat, at his desk, at the front of the class. At the end of thirty minutes, the teacher asked students to return to their seats and he proceeded to check their work individually. At this point, two of the students were randomly selected to present their work.

These observations indicated that, although students worked in groups, in some cases, the tasks (as with the one described above), were little more than worksheets that were completed with friends rather than more authentic tasks. Also as described, there were cases where the teachers set the students to work and then sat down and observed the students working. It should be noted, however, that this was not the case for the second teacher who, during group work, moved around to determine whether students were experiencing problems and to provide assistance when needed.

The interviews with each of the two case study teachers in urban schools indicated that they had purposefully changed their teaching practices to be more student-centred. One teacher commented that she found the knowledge and skills, imparted during the professional development programme, to be inspiring. To this end, the teachers stated: “The knowledge I gained during the MGMP Empowerment programme really inspired me to think about teaching English in different ways”.

Interviews with these case study teachers also indicated that they felt that, by using the new teaching strategies, the students were more enthusiastic about their study and enjoyed working with their friends. To this end, one of the teachers commented, “I am happy to see my students happy during the lessons. They work well with one another”.

The third case study teacher who was teaching in an urban school, however, made no changes to his teaching. In his opinion, student-centred approaches were ineffective and were not helpful in preparing his students for the nationwide examination. This teacher commented:

> I think that the ideas delivered in the MGMP Empowerment programme are ideal, but they are not suitable for my classroom. I always use teaching methods that meet my students' needs. You might say that I am traditional but I know what the best is for my students.

These comments were incongruent to his original interviews (held at Phase 1 of the evaluation) in which he felt that the ideas were useful and would help him to improve his knowledge and skills. This teacher found that his excessive workloads did not allow him to plan lessons that involved student-centred approaches. To this end, he commented,

> Frankly said, it is difficult to implement the ideas in my classes. The suggestions were idealistic and I cannot implement them. It is very hard to plan the lessons that were suggested in the professional development programme, it takes much time and my excessive workload does not allow me to do that.

5.3.2 Teaching Practices in Rural Schools

The rural schools in which I observed classes were all located in small villages and were between two and five hours drive away from the city. During the rainy season, one of the schools could only be reached by foot as the roads were impassable. Most of the students who attended these schools were from families whose parents worked as farmer labourers, traditional fishermen or migrant workers who lived in the city. The students whose parents worked in the cities lived with their grandparents or siblings. It was interesting to note that, during harvest season, most of the students were absent as they were needed to help their parents in the rice fields. In all three cases, the rural schools were reliant on the government funding, as parents did not pay school fees or provide donations. In each of the schools, the class sizes were larger than those in the urban schools, ranging from 40 to 46 students. The rural schools that I visited were not as well resourced. Although each of the schools had a library, they each had only one science laboratory, which

was ill-equipped when compared to the urban schools, and no language laboratory. Computers were only available to the administration staff and the principal.

In all of the classrooms in which I conducted my observations, there were 24 student desks, aligned in straight rows, and 48 chairs. Each desk had a single shelf underneath upon which books and pens were kept. There was a whiteboard in front of the classroom but no projector or screen and the teachers did not have access to a computer.

My observations of the three case study teachers in rural schools indicated that, in all cases, the pedagogy continued to be predominantly teacher-centred despite their participation in the one year professional development programme. During the 12 hours of observations, all of the teachers, on all occasions, used a format that involved a lecture followed by a question and answer session, during which the teachers asked the questions. The lessons involved the teacher explaining concepts while students copied notes from the board, after which, students completed a paper and pen exercise on their own.

The example that follows is one lesson that I observed, however it should be noted that all of my observations involved a similar format. The lesson started with the teacher reviewing the previous day's materials by asking a series of questions, after each question, the classroom was silent. The teacher was clearly disappointed with the silence and tried to make them speak by asking more questions. After some time, she asked the students to open their note books and instructed one of the students to read his notes from the day before. Reluctantly, the student read his notes in a voice that could not be heard by the rest of the class. The teacher then proceeded to revise the lesson.

Once the teacher had revised the previous day's lesson, she ordered the students to pay attention, instructing them to watch and listen carefully. The teacher wrote a grammatical formula on the whiteboard, while explaining how the formula was used and providing some examples. She spoke, for the most part, in Bahasa

Indonesia, but sometimes moved into Javanese, the local language. The teacher made frequent use of the whiteboard, writing notes that the students copied meticulously into their note books. I noticed that all of the students were silent and appeared to be listening attentively. After forty five minutes, the teacher stopped the lecture and asked the students whether they had any questions. After a few moments, during which none of the students asked her a question, she instructed them to finish copying her notes from the whiteboard.

When the students finished note taking, the teacher referred to a page in her text book and instructed the students to turn to the correct page on their book (each pair shared one text book that had been borrowed from the school library for the lesson). On the page selected by the teacher, there were exercises that the students had to complete. In one of the lessons that I observed, the teacher asked the students to work in pairs. However, the students quickly became noisy and unruly, forcing the teacher to abandon the attempt. Apart from this one attempt to get students to work together, all of the lessons were teacher-centred and followed similar format to the one described.

My observations indicated that only two of the six case study teachers had experienced any degree of success in terms of translating the ideas of the professional development programme into practice. Both of these teachers were based in urban schools and, in both cases, the use of the strategies were not always satisfactory.

5.4 Effectiveness of Professional Development in Terms of Classroom Learning Environment and Students' Attitudes (Phases 1 and 4)

To further investigate the extent to which classroom practices had changed between the beginning and the end of the programme, I collected data using two surveys, one to assess students' perceptions of their classroom experiences and another to assess their attitudes. Using the perspectives of the students was considered to be important because they are major stakeholders in the education process and because they were exposed to the learning environment on a daily basis. Past research in the field of learning environment provides strong support for students' ability to make judgements over a period of time and the usefulness of learning environment instruments for bringing together the pooled opinions of students (Fraser, 2012).

This phase of the evaluation involved examining the effectiveness of the professional development programme in terms of changes in students' attitudes and perceptions of the learning environment (between the beginning and end of the programme), the results for which are presented in Section 5.4.1. In addition, Section 5.4.2 examines whether the professional development programme was differentially effective (as indicated in the last session) for urban and rural schools.

5.4.1 Changes in Students' Perception of the Learning Environment and Attitude

To examine students' perceptions of their learning experiences and their attitudes, the Indonesian version of WIHIC and an attitude scale (described in Chapter 3) were administered to 2,417 students in 66 classes before the commencement of the professional development programme and again (to the same students), almost one year later, at the end of the programme. The results, reported in Chapter 4, provided strong support for the reliability and validity of the two instruments (the WIHIC and attitude scale), when used with this group of students.

To examine the pre–post differences, the average item mean (the scale mean divided by the number of items in that scale), for the two administrations, was calculated for each WIHIC and attitude scale. This enabled a meaningful comparison of average scores on scales with different numbers of items. Difference in the average item means are displayed in the profile for pre-test and post-test results below (Figure 5.1). The profile indicates that the post-test scores were slightly higher than the pre-test scores for all scales.

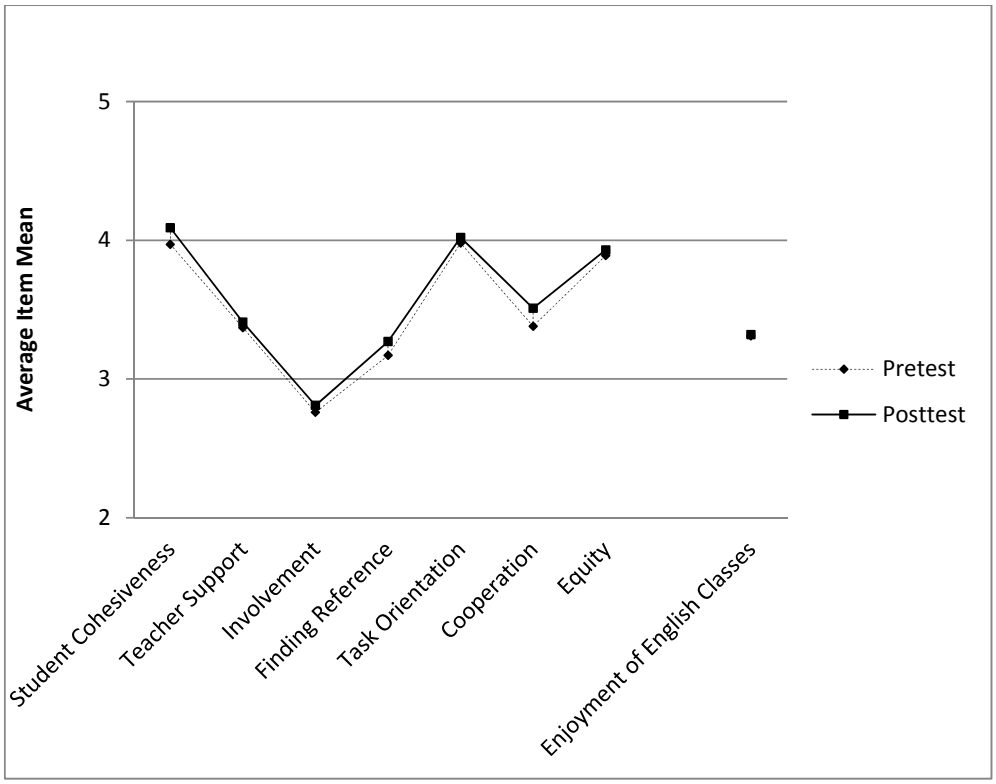


Figure 5.1 Pre-Test and Post-Test Results for Learning Environment and Attitude Scale

To further examine the pre–post differences in students' responses to the WIHIC and attitude scales, a one-way multivariate analysis of variance (MANOVA) with repeated measures was used. The dependent variables consisted of the seven WIHIC and attitude scales and the independent variable was the testing occasion (pre-test and post-test) as the repeated measure. Because the multivariate test, using Wilk's lambda criterion, yielded significant differences, for the set of

dependent variables overall the univariate one-way ANOVA was interpreted for each WIHIC and attitude scale.

In addition to the MANOVA results, which provided important information about the statistical significance of differences between the pre-test and post-test scores, it was considered meaningful to determine the magnitude of these differences and their educational importance by calculating the effect sizes (the difference between means expressed in standard deviation units). Because the effect size is independent of the sample size, it affords an additional indicator of whether the differences observed are considered educationally important or not. As Fan (2001) pointed out, statistical significance and effect size are two related sides of a coin, and they function most effectively when used together. Using the individual as the unit of analysis, effect sizes were calculated to determine the magnitude of the difference between the pre-test and post-test scores, as recommended by Thompson (2002), for each WIHIC and attitude scale.

According to the results presented in Table 5.1, there was a statistically significant (p<0.01) pre–post difference for six of the seven WIHIC scales but not for the attitude scale. All scales with a statistically significant difference showed an improvement between the pre-test and post-test: Student Cohesiveness (effect size = 0.12 standard deviations), Teacher Support (effect size = 0.03 standard deviations), Involvement (effect size = 0.04 standard deviations), Finding Reference (effect size = 0.08 standard deviations), Task Orientation (effect size = 0.07 standard deviations) and Cooperation (effect size = 0.10 standard deviations).

According to the results reported in Table 5.1, the effect sizes for the WIHIC scales with statistically significant differences ranged from only 0.04 to 0.12 standard deviations. These effect sizes all fall within the low range according to Cohen's (1992) criteria. Despite there being statistically significant differences for the six scales, the effect sizes suggest limited educational significance. It is notable, however, that the changes for all of the scales were positive, suggesting that

changes in students' perceptions, although small, were in the right direction more student-centred classrooms.

Table 5.1 Average Item Mean, Average Item Standard Deviation and Pre–Post Differences (Effect Size and MANOVA) Results

Scale	Average Item Mean		Average Item Standard Deviation		Differences	
	Pre-test	Post-test	Pre-test	Post-test	Effect Size	F
Learning Environment						
Student Cohesiveness	3.97	4.09	0.54	0.52	0.12	65.28**
Teacher Support	3.37	3.41	0.71	0.72	0.03	3.77*
Involvement	2.76	2.81	0.66	0.68	0.04	8.85**
Finding Reference	3.17	3.27	0.68	0.72	0.08	27.95**
Task Orientation	3.98	4.02	0.60	0.60	0.07	3.91*
Cooperation	3.38	3.51	0.66	0.67	0.10	45.81**
Equity	3.89	3.93	0.72	0.73	0.03	2.93
Attitude						
Enjoyment of English Classes	3.31	3.32	0.72	0.73	0.01	0.30

N= 2417 students in 66 classes

***p*<0.01 **p*<0.05

Effect size is the difference in means expressed in standard deviation units and was calculated using the formula: $d= M_1-M_2/\sqrt{[(\sigma_1^2+\sigma_2^2)/2]}$

5.4.2 Differential Effectiveness of Teacher Professional Development for Teachers in Urban and Rural Schools

Although Section 5.4.1 reported that, overall, students perceived the learning environment to be somewhat more positive after the MGMP Empowerment programme, the classroom observations (described in Section 5.3) indicated that there were differences in the extent to which the teachers in urban and rural classrooms translated the ideas from the MGMP Empowerment programme into practice. Based on these findings, the fourth research objective involved examining

whether the MGMP Empowerment programme was differentially effective for teachers in urban and rural schools, in terms of students' perceptions of their classroom learning environment and attitudes towards English classes.

Previously, Section 5.4.1 reported the use of a one-way MANOVA in exploring differences between the pre-test and post-test results for the seven WIHIC and attitude scales. In contrast, this section reports the use of a two-way multivariate analysis of variance (MANOVA) to compare the impact of the MGMP Empowerment Programme for schools in different locations (urban and rural).

For the two-way MANOVA, the independent variables were the testing occasion (a repeated measure factor involving a pre-test and post-test) and the location (rural and urban), and the dependent variables were the WIHIC and attitude scales. The sample of 2,417 students in 66 classes was again used for the analysis. Because the two-way MANOVA yielded statistically significant results overall for each of the three effects (testing occasion, location, and testing occasion x location), the univariate ANOVA was interpreted for each dependent variable for each of the three effects. Table 5.2 summarises the results. The type of effect size used for reporting the strength of association between each effect (testing occasion, location, and the interaction) for each WIHIC and attitude scale was the eta^2 statistic, which is an estimate of the proportion of variance accounted for. The results for the testing occasion (Section 5.4.2.1), location (Section 5.4.2.2) and the interaction between testing occasion and location (Section 5.4.2.3) are discussed below.

5.4.2.1 Testing Occasion

The results, reported in Table 5.2, for the two-way ANOVAs for testing occasion with control for location match the results of the previous one-way ANOVAs ignoring location (refer to Section 5.4.1). In both cases, statistically significant pre–post improvements were reported for six of the seven learning environment scales but not for the attitude scale. In all cases, the effect sizes were small, ranging from 0.01 to 0.03 for scales with a significant difference. The direction of the differences

were positive in all cases, indicating that the post-test scores were higher than the pre-test scores.

Table 5.2 Two-Way MANOVA/ANOVA Results (*F* and Eta^2) for Testing Occasion and Location for Each WIHIC and Attitude Scale

Scale	Two-Way ANOVA Results					
	Testing Occasion		Location		Testing Occasion x Location	
	F	Eta^2	*F*	Eta^2	*F*	Eta^2
Learning Environment						
Student Cohesiveness	80.43**	0.03	113.85**	0.03	21.19**	0.01
Teacher Support	7.44*	0.02	95.57**	0.02	15.03**	0.01
Involvement	9.92**	0.02	23.26**	0.01	1.25	0.00
Finding Reference	34.73**	0.03	51.71**	0.02	11.01**	0.01
Task Orientation	6.26*	0.01	46.21**	0.02	7.17**	0.01
Cooperation	58.13**	0.03	51.71**	0.02	21.09**	0.01
Equity	5.75	0.01	78.45**	0.02	11.60**	0.02
Attitude						
Enjoyment of English Classes	3.17	0.00	9.53**	0.00	35.67**	0.01

N= 2417 students in 66 classes present for both pre-test and post-test

**p*<0.05 ***p*<0.01

5.4.2.2 Location

Table 5.2 reports ANOVA results for whether differences exist between for students in rural and urban locations, regardless of the testing occasion. As shown in Table 5.2, statistically significant ($p<0.01$) differences were found for all seven WIHIC and attitude scales. The proportion of variance was small and ranged from 0.01 to 0.03. In all cases, the direction of the differences were positive, indicating higher scores for student in urban schools.

Because significant interactions exist between the testing occasion and location for six of the seven WIHIC scales and attitude scales, as shown in Table 5.2, these interactions are discussed separately in Section 5.4.2.3.

5.4.2.3 Interactions between Testing Occasion and Location

Information about the differential effectiveness of the MGMP Empowerment programme for teachers in urban and rural schools was obtained by examining the interactions between testing occasion and location identified through the two-way ANOVAs. The results in Table 5.2 show that a statistically significant interaction between testing occasion and location existed for six of the seven learning environment scales (with the exception being involvement) and the attitude scale. Therefore the independent interpretations of pre–post differences and location differences are valid only for the Involvement scale.

Table 5.2 shows that the amount of variance accounted for by the statistically significant (p<0.01) interactions between testing occasion and location (in terms of the eta^2 statistic) was 0.01 for the scales Student Cohesiveness, Teacher Support, Finding Reference, Task Orientation, Cooperation and Enjoyment of English Classes and 0.02 for Equity. To better understand the interactions between testing occasion and location, Table 5.3 shows the average item mean for pre-test and post-test, average item standard deviation for pre-test and post-test, and difference between pre-test and post-test for each learning environment and student attitude scale. The effect size for each is reported in the last column.

Pre–Post Differences for Learning Environment Scales

Examination of the average item means, separately for urban and rural students, in Table 5.3, reflects the direction of the urban-rural differences, with urban students' scores being higher for all scales of the learning environment scales for both the pre-test and the post-test. These differences for urban and rural students are displayed graphically in Figure 5.2.

Table 5.3 Average Item Mean, Average Item Standard Deviation and Pre–Post Differences (Effect Size) for Urban and Rural Schools for Learning Environment and Attitude Scales

Scale	School Region	Average Item Mean		Average Item Standard Deviation		Difference Between Pre-test and Post-test
		Pre-test	Post-test	Pre-test	Post-test	Effect Size
Learning Environment						
Student Cohesiveness	Urban	4.02	4.23	0.50	0.47	0.21
	Rural	3.93	3.99	0.56	0.53	0.06
Teacher Support	Urban	3.44	3.58	0.68	0.74	0.10
	Rural	3.32	3.30	0.72	0.69	-0.02
Involvement	Urban	2.80	2.88	0.61	0.68	0.06
	Rural	2.73	2.77	0.69	0.67	0.03
Finding Reference	Urban	3.22	3.40	0.64	0.71	0.14
	Rural	3.14	3.19	0.7	0.71	0.04
Task Orientation	Urban	4.03	4.12	0.57	0.56	0.08
	Rural	3.96	3.95	0.61	0.61	-0.00
Cooperation	Urban	3.41	3.65	0.63	0.64	0.18
	Rural	3.36	3.42	0.68	0.66	0.04
Equity	Urban	3.96	4.09	0.71	0.71	0.09
	Rural	3.85	3.83	0.72	0.72	-0.01
Attitude						
Enjoyment of English Classes	Urban	3.19	3.36	0.73	0.77	0.11
	Rural	3.39	3.30	0.71	0.71	-0.06

N= 2417 students in 66 classes present for both pre-test and post-test

*p<0.05 **p<0.01

Effect size was calculated using the following formula: $d= M_1-M_2/V[(\sigma_1^2+\sigma_2^2)/2]$

It is interesting to note that, although the combined sample of urban and rural students showed a significant pre–post improvement in terms of each of the WIHIC scales, there was a statistically significant interaction (reported in Table 5.2) for these two groups, for six of the seven WIHIC scales (the exceptions being the Involvement scale). This suggests that a valid interpretation for these scales could be obtained by examining the results for urban and rural classes separately. Figures 5.3 to 5.8 depict the six significant interactions for the learning environment scales.

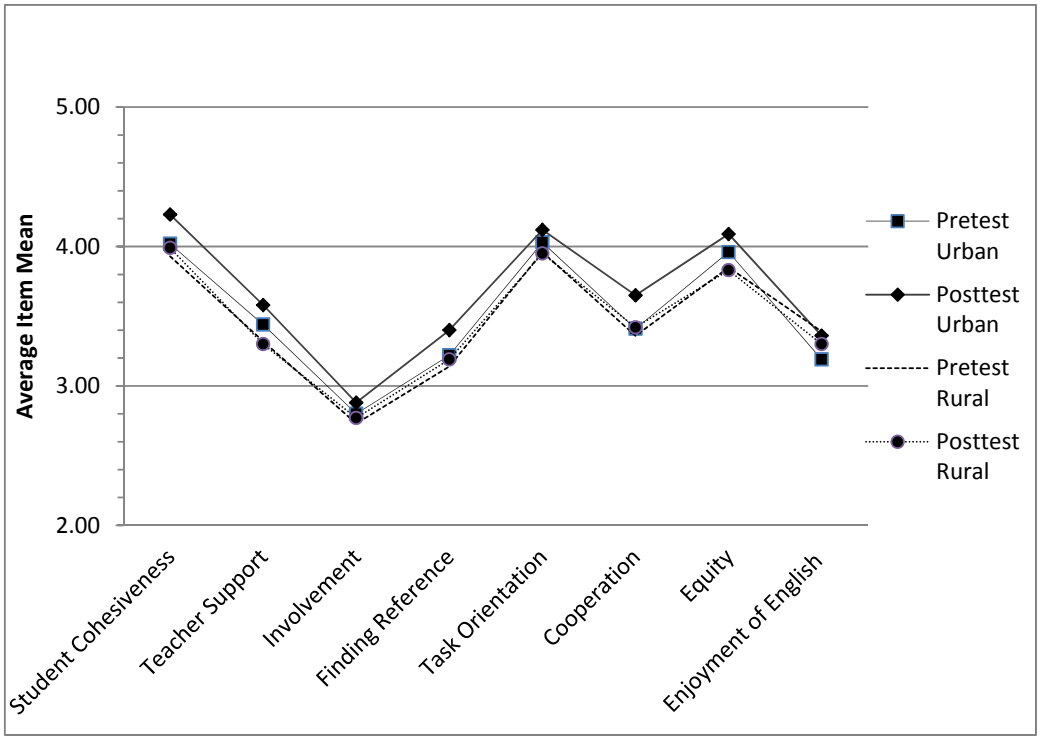


Figure 5.2 Pre-test and Post-test Scores for Students' Responses to WIHIC and Attitude Scales for Urban and Rural Students

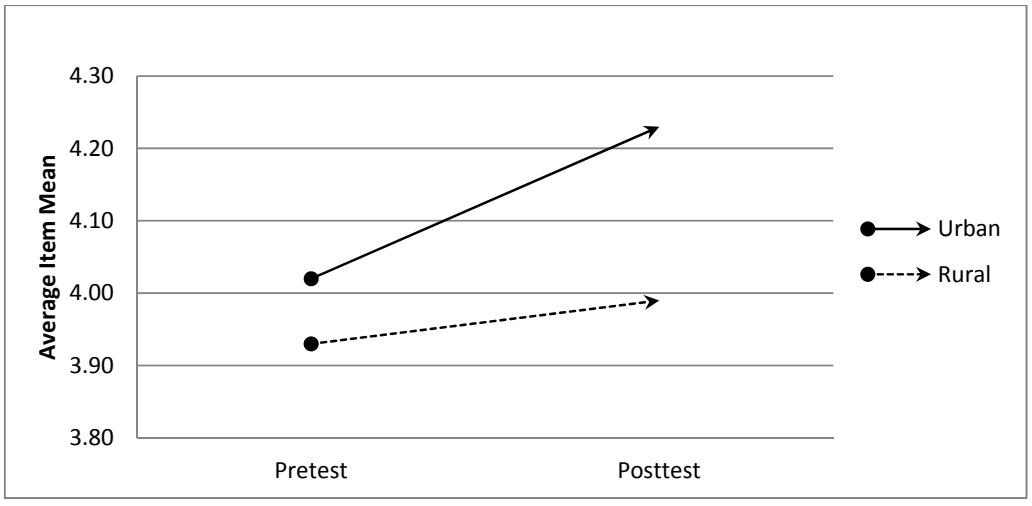


Figure 5.3 Interaction Between Testing Occasion and Location for Student Cohesiveness

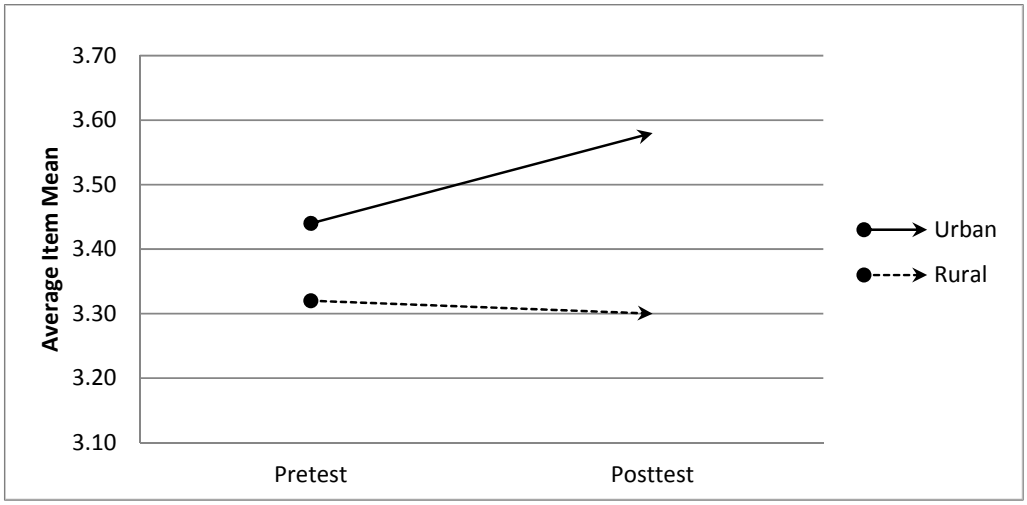


Figure 5.4 Interaction Between Testing Occasion and Location for Teacher Support

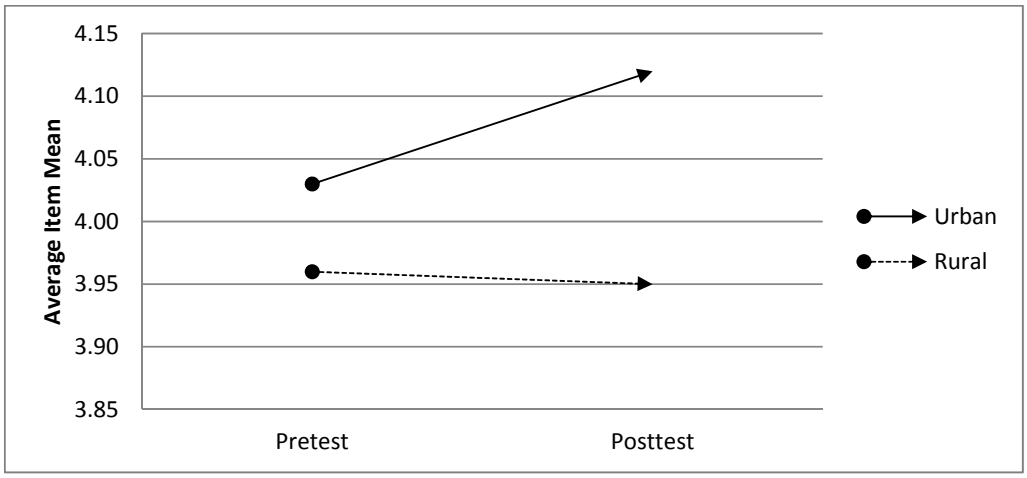


Figure 5.5 Interaction Between Testing Occasion and Location for Task Orientation

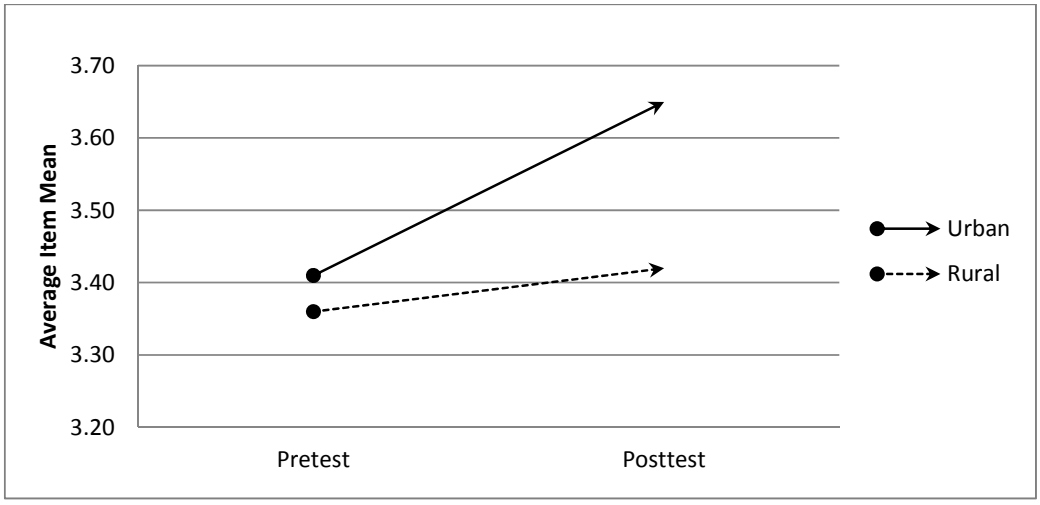


Figure 5.6 Interaction Between Testing Occasion and Location for Cooperation

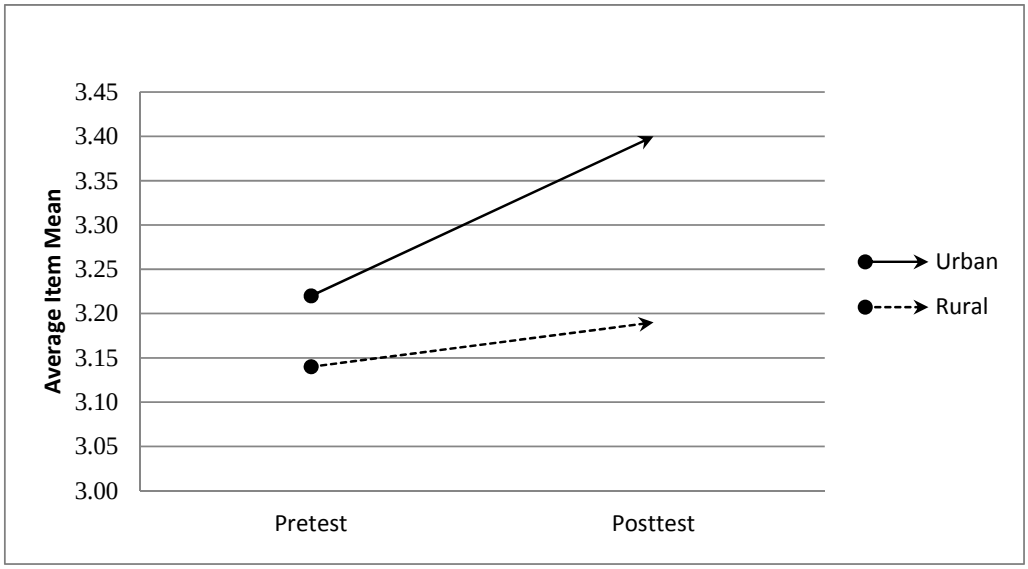


Figure 5.7 Interaction Between Testing Occasion and Location for Finding Reference

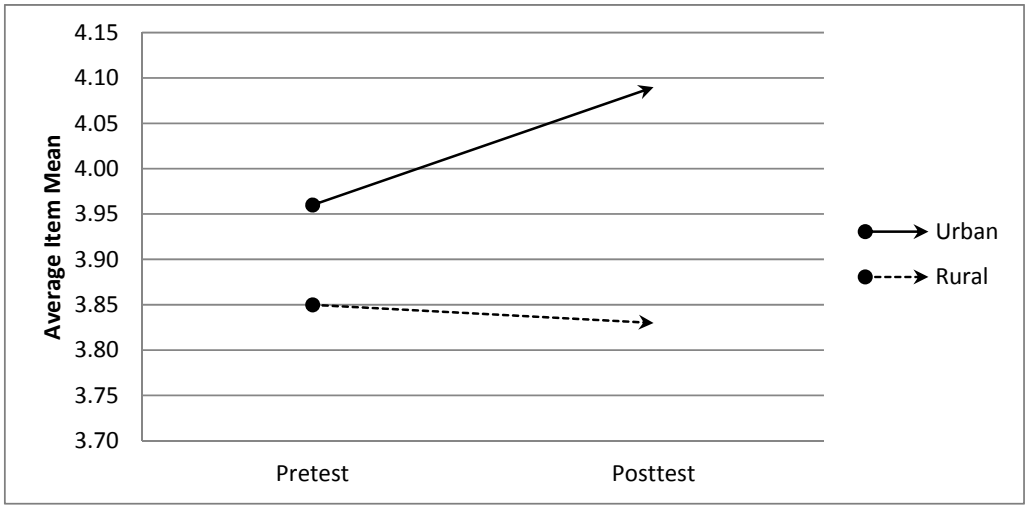


Figure 5.8 Interaction Between Testing Occasion and Location for Equity

Each of the interactions depicted in Figures 5.3 to 5.8 shows a locality gap at the pre-test which is larger at the post-test. For all scales, the pre–post change for the urban group is greater and positive in direction whereas the pre-post changes for the rural group are smaller and, for three of the seven scales, negative in direction. In all cases, the effect size for the pre–post differences in students' perceptions of the learning environment was larger for urban classes (ranging from 0.08 to 0.21 standard deviations) than for rural classes (ranging from 0.00 to 0.06 standard deviations) (see Table 5.3).

Interviews with students were used to help to clarify student responses and to explain the differences in pre–post changes. Interviews with students in urban and rural schools indicated that the students in different school localities had interpreted the questionnaire items in similar ways and, given this, investigating the reasons for this differential effectiveness became more imperative. Therefore, classroom observations and interviews with students helped me to make sense of the questionnaire results in both urban and rural schools.

Classroom observations and interviews with students suggested that the differences in the pre–post changes in students' perceptions for the four scales of Student Cohesiveness, Teacher Support, Task Orientation, and Cooperation might be attributed to the introduction of more student-centred strategies in the classes of urban schools. As discussed in Section 5.3, two of the three case study teachers in urban schools had attempted to change their teaching approaches (see Section 5.3). In these classes I observed that these teachers provided opportunities for the students to work in groups. It was important to note that interviews with students suggested that these group exercises also were carried out when the researcher was not present. During my observations, I noted that, during group work, the students generally worked well together and that they were helpful and supportive of one another. To this end, a student from one of the two urban schools stated:

> Sometimes my teacher gives us the group tasks and, usually, she selects the members of the groups. I usually work well with my friends. But sometimes, when I work with lazy students, there is not much team work.

Another student commented:

> My teacher usually gives us group tasks. The students choose the members of the group. I usually choose my diligent and smart friends. Working with smart friends is interesting because I can work well with them. Everybody gets a similar opportunity to do the tasks.

The students in the classes of these two teachers expressed their desire to work in groups. To this end, one student commented, “When I work with my group, I get along well with them. I like it”. The students whom were interviewed also indicated that they felt that working together was valuable. They articulated that they worked better and learned more when they worked in group. To this end one student commented, “Working with friends is interesting because I can learn from them – especially when I do not understand the lesson”. These students also commented that, when they worked with others, they were able to get to know one another. One student commented, “If I work with my friends, I make friends and this is interesting for me to know more about them”.

The students whom were interviewed also felt that, during group work, their teachers were generally helpful and that they received support from them. One of the students stated:

> I like my English teacher, he is very interesting. Sometimes, he tells jokes which makes our English classes less boring. He never gets angry easily. He always attends to us when we have problems when working on our tasks.

On the contrary, all of the observations that I made in rural classrooms, during the one year of the MGMP Empowerment programme (Section 5.3), provided no evidence of group work or any other student-centred activity. Typically, the classrooms were arranged with desks in rows and with the students facing their teacher at the front of the class (as described in Section 5.3.2). During my observations the students worked independently with no opportunities for interaction. Frequently, the students were told to copy notes from the blackboard and there were, in general, few opportunities for discussions or questions. Of all of my observations in rural schools, only one attempt was made by one of the teachers to ask their students to work in pairs. As described earlier, the session was unsuccessful as students became loud and unruly, forcing the teacher to abandon the attempt. It would appear that the students did not understand what was expected of them and became off-task. Interviews with students indicated that the

teacher did not attempt any pair or group work when the researcher was not present.

When asked, the students in rural schools stated that they preferred to work independently. Students whom were interviewed all expressed that they were unlikely to help each other and did not support one another because they felt that, they themselves, experienced difficulties in learning English. For example, one student from a rural school commented:

> I do not help my friends who are having problems with their work because I have similar problems. Also, girls work hard and of course girls will only help girlfriends. The boys in my classroom are lazy and I can't work with lazy boys.

Similarly, another student from a rural school commented:

> I cannot help my friends because I have the same problems. Sometimes, I am not sure about my ability, so I just tell them I don't know the answer.

Besides experiencing difficulties in learning English, these students also felt that they were not given opportunities to work in groups and, as a result, had fewer opportunities to collaborate, work together or to each help each other. To this end one student from a rural school stated, "My teacher rarely asks us to work in groups. The teacher always asks questions after the lessons. She asks students to answer them one by one".

Whilst it is acknowledged that there may be other factors that might contribute to the differences between the changes experienced by students in rural and urban schools, the observations and interviews with students suggest that the use of group work might explain the changes for these four scales.

The Finding Reference scale describes the extent to which skills and processes of finding references and their uses in problem solving and investigations are

emphasised. Classroom observations and interviews with students indicated that the extent to which the schools were resourced and the economic disparities between the students attending rural and urban schools may have contributed to the differences in changes to this scale.

During my observations, it was clear that the students in urban schools used references more regularly than their counterparts in rural schools. When the students in urban schools were working in groups they often used reference books or the internet to support their work. When interviewing these students it was apparent that they felt positive about using a range of references. To this end a student from an urban school stated, “My school provides every student with an English textbook but I always make use of my English reference books to support my works”. Another student commented that using reference books and internet helped him to do his homework:

> Besides having the text book that the school lends me, I also have some reference books that my mother bought me. I use these books as the resources when doing my homework. If I cannot find the answer in my reference books, I usually browse the internet to find the answers.

In contrast, students in rural schools did not have reference books available to them. Although the school provided a text book, these were borrowed from the library for the duration of the English class and, when the class was finished, returned so that they could be used by other classes. Interviews with students in rural schools indicated an additional disadvantage of not having access to other resources because of their financial situation. As stated by one of the student from a rural school:

> I do not have any English books. The school provides only a few English text books in the library and the students can only use the books during the English classes. The students never bring the books home. If there is homework, the students must copy the book by writing down the homework. My parents

cannot buy reference books for me. I know that they do not have the money to afford it.

It would appear that the small changes in pre-post perceptions of the use of references experienced by students in rural areas may not be because teachers are unprepared to use a range of references, but rather, that they do not have this resource available to them.

The Equity scale assesses the extent to which students perceive that they are treated equally or fairly by their teacher. As with the other learning environment scales, the locality gap existed at the pre-test but was larger at the post-test. Observations coupled with interviews with students, helped to explain these differences in pre–post changes. It was interesting to note that, throughout the lesson, the teachers in urban classes tended to provide opportunities to all of their students to ask and answer questions. Students in these classes expressed that, during the interviews, they felt that the teacher was fair in terms of the amount of attention that the teacher paid them and the opportunities afforded to them in a range of situations. To this end, a student from an urban school claimed, “Most of the time, my teacher asks questions to the class and we get the same opportunities to answer her questions”. Another student commented, “Some of my classmates need more help from the teacher and the teacher will help them”.

In contrast, my observations in the classes in rural schools indicated that the teachers may not provide the same opportunities to all of their students. My observations indicated a strong tendency for the teachers in rural schools to provide more assistance and opportunities to students who were considered to be high achievers. Interviews with the students in rural schools supported these observations, indicating that the ‘bright’ students took up much of the teacher’s time and attention, leaving the remaining of the students to their own devices. To this end, one of the students from a rural school commented:

> The teacher does not like lazy and naughty students. She never attends to them or helps them or checks their work. Personally, I never ask my teacher questions because I do not know what to ask. My friends also never do that, either.

It would appear that the blatant disparities between the attention given to students who were capable and those who were not, was reflected in the pre–post changes. This point is explored further, from the teachers' view-point in Section 5.5.1 .

Pre–Post Differences for Enjoyment of English Classes

When the sample was combined, the results suggested that the pre–post differences for the students' enjoyment of English classes were not statistically significant, however, the results reported in Table 5.3 show that there was a statistically significant interaction for this scale which is depicted graphically below in Figure 5.9. The results show that, in urban schools, students' enjoyment of English classes improved (effect size = 0.11 standard deviations), whereas the enjoyment of English classes for their rural counterparts declined marginally (effect size = 0.06 standard deviations).

The analysis of data collected during classroom observations and interviews suggested that the differences between urban and rural students' enjoyment of English classes may have been influenced by the teachers' manner taught and the materials presented during the lessons. Interviews with the students from urban schools suggested that they perceived their teachers to be friendly and that they provided an environment which made learning English more enjoyable. A student from an urban school who was interviewed commented:

> I always feel when I came to English classes. English is always interesting and my teacher is very friendly; she does not get angry unexpectedly. She gives opportunities for students to develop their thinking and she likes telling jokes.

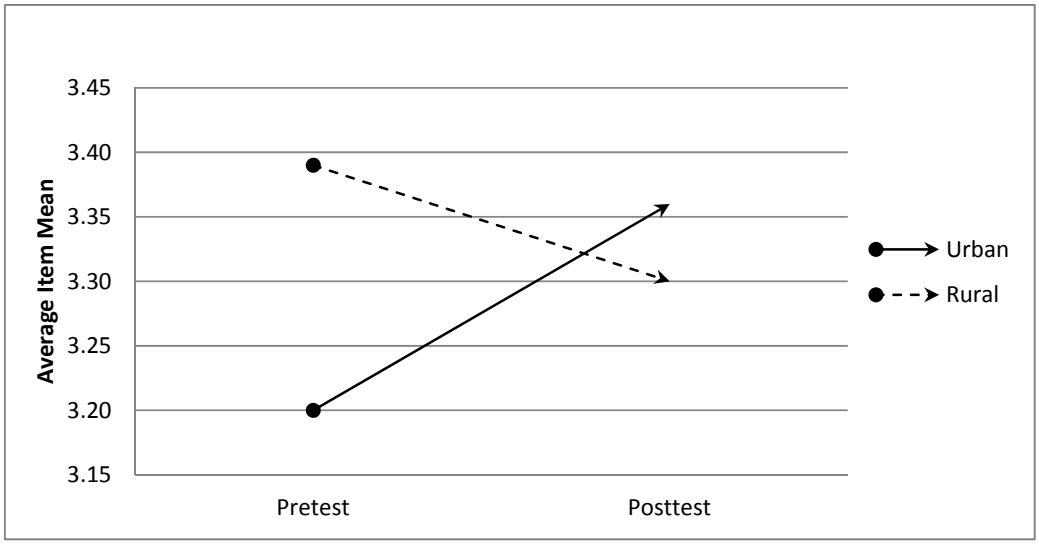


Figure 5.9 Interaction between Testing for Occasion and Location for Enjoyment of English Classes

Observations in the classes of the two case study teachers who changed their practice (both of whom were located in urban schools) indicated that the students enjoyed the opportunity to get involved and to interact with each other during the learning activities. A student from an urban school stated:

> I feel very happy and have a great time during English lessons. My teacher always teaches interesting topics. The English materials are interesting. Sometimes I learn about cultures or cultural events of other people living in the other parts of the world. I also enjoy learning the places of interest in the other countries.

Besides their positive opinions of the interesting materials presented by their teachers, the interviews also suggested that the students from urban schools viewed learning English as worthwhile and important for their future. These students commented that they needed to master English for their future, for example:

> I want to be a successful international businessman. To reach my ideal, I must master English. If I do business abroad, I will need to speak English to be able to negotiate with other business people. Also, when I want to know about the

> international market, I must understand the news about the market that is happening in the world. That is why I think English is very important for me.

On the other hand, the students from rural schools whom were interviewed were more likely to view English as difficult and boring to learn. In addition, these students did not appear to like the way that the teacher taught. One student commented:

> For me, English is very difficult. I find it difficult to understand, especially the English grammar and vocabulary. It is boring. I like drawing better than doing English tasks. If the teacher finds me drawing during English class, she gets angry with me.

These interviews also suggested that students from rural schools did not view English as either important or useful to their future. As one student commented:

> I know that English is an international language, but I don't think that I need to master it. After graduating from this school I am not sure that I will continue my studies at a higher level. My parents do not have money to support me so I think I will work as a labourer in a sock company. In fact I would like to continue my studies at a vocational school but I cannot. I have two younger brothers who also need to be supported. That is why English is not important to me, because I will not use it again after finishing junior high school.

Based on the interviews with the students, it would appear that that the students in urban schools appeared to enjoy their English classes more than their counterparts in rural schools because the teachers were friendlier and more lenient and they were more likely to view learning English as valuable to their future.

5.5 Contextual Factors that Promote or Impede the Programme Implementation (Phase 5)

Teacher professional development is a systemic process that considers change over an extended period of time and takes into account all levels of organisation (Guskey, 2000, 2002; Johnson & Johnson, 1995; Sparks, 2002). Without a systemic approach, organisational variables can hinder or prevent the success of improvement efforts, even when the individual aspects of professional development are done right. Spark and Hirsh (2002) added that, when viewed systematically, professional development was seen not just in terms of individual improvement, but also in terms of improvements in the capacity of the organisation to solve problems and renew itself.

In the present study, classroom observations and interviews suggested that there were constraints which may have prevented the teachers from implementing the ideas presented during the professional development programme. Therefore, Phase 5 of the evaluation examined contextual factors that might promote or impede the translation of the professional development programme into practice. Classroom observations and interviews with the six case study teachers suggested that there were three overarching contextual factors that were influential, these being: the instructional context (discussed in Section 5.5.1), the organisational support (discussed in Section 5.5.2); and system-wide policy (discussed in Section 5.5.3).

5.5.1 Instructional Context

Analysis of the data indicated that the instructional context in which the teachers worked made a difference to the extent to which they were both able and willing to translate the ideas promoted during the professional development into practice. For these teachers, the factors that appeared to be most influential were the socioeconomic background and behaviours of their students, the class size and the availability of resources. It would appear that, despite the teachers' best intentions

to implement the new ideas, as indicated in Section 5.2, these factors impacted on their ability to do so.

Interviews with the two case study teachers in urban schools (both of whom had made changes to their teaching practice) indicated that their willingness to implement the ideas was due, largely, to the status of the school (which attracted students who were achieving well academically) and the socio-economic status of the families of the students. There were, according to these teachers, several benefits to being in such a school. They felt that their students did not generally experience difficulty in learning English and that they were, for the most part, motivated to learn and to study hard. The two teachers both agreed that, generally speaking, the students in their schools were polite and well behaved. It would appear that the combination of these factors made it easier for them to try out the new ideas. To this end, one teacher stated:

> I have no problems in using the new teaching methods. Most of my students are 'straight A' students and they are from families of middle to high economic and social status background. I am very happy with the fact that the students are well behaved and their parents are very supportive [financially]. The parents consistently encourage and motivate their children to study hard.

The other teacher commented in a similar vein,

> My students are nice; they have high academic standards and come from the families of high economic background. You know how high standard students behave; they are polite and nice to everybody. Some students are naughty but the teachers can control them easily. The supports from the parents and society are good, too. These points mean that I do not have problems teaching with new approaches.

In stark contrast, the three teachers from rural schools were all of the opinion that most of their students were lazy, disruptive and unmanageable. Based on my interviews with these teachers (coupled with the classroom observations), it would

appear that issues related to discipline took up much of the class time. One of the teachers complained that maintaining discipline was a problem because they no longer had the liberty to physically punish students. The interviews with the three teachers in rural schools made clear that the students' lack of motivation and problems related to discipline was a pressing issue. To this end, one of the teachers stated:

> It is very boring to teach lazy students in a low standard school. I feel it is far too much effort to change. If you were me, dealing with lazy students all the time, what would you do? Punishment does not work anymore. The students feel fine with any kind of punishment given by the teachers. They don't feel guilty. Students that make noise in the class and play truant are everyday problems for me.

All three of the case study teachers felt that, by using teacher-centred methods, they were better able to keep the class quiet and under control. During my interviews with these teachers, it became clear that, perhaps much of their hesitancy to adopt the student-centred techniques stemmed from their fear of disruption and losing control of the class.

Unlike the two case study teachers who were teaching in urban schools, who felt that their students were achieving well and were not struggling, the teachers who were teaching in rural schools felt that their students were not motivated and were low achievers. The teachers in rural schools argued that, because of this, their students would not benefit from student-centred methods. To this end, one of the teachers commented:

> I realise that I still teach English in a traditional way. I do not use any media whatsoever when I teach because I think it's useless. My students are villagers; they are fine with the way I am teaching. If I use authentic language resources, as suggested in the professional development programme, I don't think they would understand. That's why I believe that these students learn more effectively if I teach in the traditional way.

One of the three teachers who were teaching in a rural school recalled that, when the professional development programme first started, he was very enthusiastic about the ideas and implementing them in his classroom. However, he quickly realised that the ideas were not applicable to his teaching situation.

> I thought that the suggestions that were delivered in the professional development programme were very good but when I came back to my school, I found that those suggestions could not be implemented in my classrooms. The conditions of my students, their low academic ability and lack of motivation to learn hindered the implementation.

Of concern was the confession of one of the teachers who was interested only in students who were engaged during the learning process. This teacher explained that, because of the time constraints, focusing on students who were doing well academically and were keen to learn was more important than paying attention to the students who were off the task. To this end, the teacher commented:

> I have to cover all of the materials in a limited time. There is no time to take care of lazy and naughty students. My students and I have agreed to a rule; those who do not want to learn can stay in class but they have to be silent or they can stay outside. It's hard to manage 46 students and to ensure that they are always focused in my lessons. Therefore, I only concentrate on the students who want to learn.

Further to this teacher's comment, it would appear that the number of students in the classrooms may also have influenced the teachers' ability to implement new ideas. The class sizes in urban schools were considerably smaller (averaging around 20 to 30 students) than classes in rural schools (generally more than 40 students). It would appear that the smaller class size, coupled with a lack of discipline problems made implementing the ideas easier in urban schools. All three of the case study teachers in the rural schools, on the other hand, felt that the large class sizes hindered their attempts to provide more student centred activities, such as group work. One of the case study teacher from a rural school commented:

> It is not easy to control 46 students in one classroom and to get them to work in groups. [When I tried to use group work] the students did not get along. My students like working by themselves. Working in groups slowed them down because they did not work well with each other.

Although all three of the teachers in rural schools claimed to have tried to use some of the suggestions, they all felt that these efforts were unsuccessful. These teachers felt that the new teaching approaches, particularly those that encouraged students to be cooperative, were not effective in improving the students' performance in English. Their belief that the strategies were not suitable and that the students were unlikely to benefit from working in groups proved to be a significant hindrance to implementing the professional development ideas.

Finally, it would appear that access to suitable resources could have hindered the ability of teachers to implement the ideas suggested in the professional development programme. None of the teachers in urban schools mentioned the issue of resources, or the lack thereof. In contrast, all three of the case study teachers from rural schools expressed that their schools lacked the resources and facilities required to support the activities suggested during the professional development programme. One of the case study teachers from a rural school commented, "My problem is that, if I want to implement a new technique in my classroom, I have no learning aid materials that are suitable". Another teacher from a rural school stated that:

> My problem is the lack of facilities, such as a language laboratory or a specialised English classroom. I know that some of my friends use a laptop, but I do not think it is a real change. The laptop is only used to replace the blackboard. Rather than write on the board, they type the materials. Just that. It is not used as a new medium of teaching, it's only, you know, replacing the chalk and the board.

It would appear that the instructional context was likely to promote or inhibit the successful implementation of the MGMP Empowerment programme. In particular,

the findings suggest that the behaviour of the students, class size and access to resources (all of which could be influenced by the socioeconomic status of the families attending the school) were factors that influenced the ability and willingness of the teachers to implement the new teaching strategies.

5.5.2 Organisational Support

Within the context of this study the organisational support, and how this might impede or promote the implementation of new ideas, was focused on understanding the support from the principals and higher level of administration within the education system. Interviews with the case study teachers, from both urban and rural areas, indicated that the support that they received from their principals was limited to 'releasing' them from teaching duties so that they were able to participate in the programme. In all cases, the teachers indicated that the school's principal did not support their efforts as they tried to change. To this end, one of the teachers commented:

> In my case, there is no sense that my principal cares for her teachers. The principal does not pay attention to the teachers who do their job well. She does not treat the lazy teachers differently. There has been no support from the principal with the implementation of the innovations. One day, when I asked for reference books for my students, my request was rejected.

The interviews with the six case study teachers indicated that, in all cases, the teachers felt that the principal did not provide opportunities for them to work collaboratively, to help each other, or to plan for change. For example, the teachers' from urban schools intended to conduct lesson study sessions (peer planning and peer observation) in their schools but they were hindered by their teaching schedules.

> I invited my colleagues to implement one of the suggestions in the professional development (lesson study as an approach of collaborative professional

> development), but it did not work. Our teaching schedules did not match. If the principal could have helped us by scheduling policy to provide opportunities for peer planning and peer observations, it would be helpful.

A teacher from a rural school also commented:

> It's difficult, you know. The teachers have excessive workloads and it's not easy to find knowledgeable others [resource person]. Funding is a must and it's not easy to get the funding from school. There are a lot of problems that make the implementation of professional development difficult if the teacher does not have the support from the school.

The teachers from rural schools claimed that the principals did not create a culture that encouraged change. For example, in one school the discipline policy was not strongly enforced by the principal, which indirectly hampered the teacher's efforts to implement the professional development ideas. This teacher commented:

> In my school, discipline is a major problem. Not only the students but also the staff come to school late. How can a school that starts at 7 a.m. and still function when only 5 % of the school staff arrive on time? If the discipline of the school staff is a problem, then it is impossible to consider school improvement or school reform. If the principal and the teachers are to be models for the students then surely the staff should arrive at school on time before they can punish the students.

In addition to a lack of principal support, all six of the case study teachers made clear that there was no practical support from their district superintendent. One of the case study teacher commented, "The superintendent does not have time to supervise the implementation of the professional development ideas. He is very busy, and he does not have any knowledge about the innovation, either".

Interviews with the case study teachers indicated that neither their principals nor their superintendents seemed to be aware of what they were doing or what the

professional development programme involved. None of the case study teachers were observed by either the principals or the superintendents during the implementation; consequently, they did receive any encouragement or constructive feedback which they felt could have helped them to improve their teaching. One of the teachers stated, “After the professional development programme, there are no follow up activities. Neither the principal nor the superintendent observed whether we were doing the right things or not”.

Based on the interviews with the case study teachers, it would appear that the teachers felt that the organisational support was inadequate, thereby impeding the success of the professional development programme.

5.5.3 Nation-Wide Examination

Interviews with the six case study teachers strongly suggested that the nationwide, leaving examination may have been counterproductive to the effective implementation of the programme ideas. All of the case study teachers made clear the importance placed on the high-stakes nationwide leaving examination (Ujian Nasional), the scores for which, determine whether students are retained at the same level for another year or accepted in to a senior high school (the higher level). Those students with higher scores were more likely to be accepted in to a ‘star’ senior high school (a school with outstanding results as measured by the number of students who enter a ‘star’ university). Despite the fact that, in 2006, the government introduced an outcome-focused curriculum, the scores on the leaving examination were still considered to be the highest priority.

All six of the case study teachers indicated that, given the importance of the grades that students achieve in the leaving examination, they were unlikely to change their teaching practice and would remain focused on providing the content that the students required to pass. In particular, each of the case study teachers from rural schools claimed that, on a number of occasions, the competitive nature of examination-driven curriculum put pressure on them to stick to what they knew

and prevented them from experimenting with unfamiliar methods. They claimed that, to experiment with new methods, would take up precious time and would have unknown results. All of the case study teachers viewed the traditional approach as more efficient for transmitting the information required for the examination and allowed them to prepare their students more effectively. A teacher from a rural school explained his anxiety:

> I have not changed the way that I teach; it is still traditional and teacher-centred. I realise that I have not changed my way of teaching, as suggested in the teacher professional development, but the nationwide examination at the end of the year is more important. My target is to prepare my students to sit that exam. If my students fail the exam, it would present a serious problem for me. The failure of my students in the exam would be bad for my career and reputation. I do not want this happen to me. I cannot use the techniques or activities in my classes because I must teach to the exam.

It was interesting to note that, despite the changes that occurred in the classes of the teachers in the two urban schools; these teachers were far from convinced that student-centred activities were useful in preparing students for the nationwide examination. One of these case study teachers commented,

> When I use the student-centred approach, I am concerned that I am not preparing the students for the nation-wide examination. This presents a problem because although teachers believe that the student-centred approach is effective in developing problem-solving skills, it is not effective in helping students to prepare for the examination.

These revelations were somewhat incongruent to the observations that showed two teachers in urban schools to be implementing changes. Further interviews with these two teachers suggested that they were willing to try out the new ideas, not because they felt that they would be useful, but because almost all of their students attended private after-school tuition classes. These teachers felt that if the student-

centred approaches failed to prepare the students for the examination, the private tuition would serve to fill the gap. To this end, one of the teachers said:

> Almost all of my students attend private English tuitions after school. This is very advantageous for me because they have another place where they can learn English. Therefore, the students do not fully rely on me to learn English, which is very good.

The other teacher commented:

> I am happy that my students attend private English tuitions. There they learn the factual knowledge that they need to pass the national examination. It helps me sometimes.

These interviews with teachers helped to explain why the two case study teachers in urban schools were willing to try the student-centred strategies. These findings were consistent with expectancy theory, that is, an educational innovation is more likely to be implemented if the innovation valuable and the likelihood of its success are high (Fullan, 2007; Guskey 2002; Shah & Higgins, 1998).

The third teacher from an urban school, who continued to use a teacher-centred approach, felt that the new curriculum was too complicated. To this end this teacher stated:

> The new English curriculum is complicated. The new content is too advanced for students at the lower high school level. As I am teaching year 9 students, I need to focus on preparing them for the national leaving examination. I have to manage my limited time so that I can prepare the students for the examination. I do not want to take any risks because my students must perform well.

This teacher went on to criticise the national final examination:

> I think that the government should change the policy of the national final examination. The government must be made to see the implications of the national examination policy at the school level. The government needs to ask the teachers about their burdens and the influences of the policy on the teaching and learning processes at the classroom level. Because of this policy, the education stakeholders [parents, society and others] are too interested in the examination results; they do not pay attention to the learning and teaching processes. I think the national final examination is fine, if it only functions as a national measurement to get information about the quality of education. I do not agree that the results of the examination should determine whether students fail or succeed at the lower high school level. It is not fair that the success of a student is determined by the results of a four day test.

The preliminary interviews, conducted at Phase 2, during the initial stages of the professional development programme (reported in Section 5.2), indicated that all six of the case study teachers were excited about the new ideas suggested during the programme and were keen to use them in their classrooms. In reality, however, a range of factors made it difficult, particularly for teachers in rural schools, to implement them. It would appear that the teachers in rural schools quickly became reluctant to try new strategies and vigorously defended their teacher-centred approaches, claiming that these methods worked best for the students they were teaching. The findings my study reflect past research which suggests that contextual factors can lead to a resistance adopting the new techniques (Brodie, Lelliott & Davis, 2002; Loucks-Horsley & Matsumoto, 1999).

5.6 Chapter Summary

This chapter reported the evaluation of the MGMP Empowerment programme. To examine the effectiveness of the teacher professional development programme, data was collected at five points to coincide with the five phases of the evaluation model to examine: 1) the teachers' views of the relevance and utility of the

professional development programme, 2) the teachers' use of knowledge and skills, 3) changes in students' perceptions of their learning experiences and attitudes, and 4) contextual factors that promoted or impeded the implementation.

Information gathered from teachers' reflective journals and interviews were used to evaluate the teachers' views of the relevance and utility of the teacher professional development programme (Phase 2). The findings indicated that the teachers generally perceived the professional development experience positively, however, they were critical about two aspects of the delivery, these being the calibre of two of the four trainers and, the tight time schedule required to complete the programme.

Data gathered using classroom observations and interviews were used to examine the teachers' use of knowledge and skills (Phases 1 and 3). A minimum of four hours of observations (one hour of which was conducted prior the commencement of professional development and the remainder near the end of the programme) were conducted in each of the classes of the case study teachers. The findings indicated that there were differences in how the teachers translated the ideas, imparted during the professional development programme, into their classroom practices. Two of the case study teachers (both of whom were teaching in urban schools) attempted to change their teaching practices to be more student-centred. These teachers used a variety of teaching approaches to encourage their students to work in groups. However the remaining case study teachers (one teaching in an urban school and three teaching in rural schools) did not show any signs of changing their teaching practices and all of whom continued to use predominantly teacher-centred methods.

To examine changes in students' perceptions of their learning experiences and attitudes (Phases 1 and 4), the study involved a pre–post administration two surveys: the What Is Happening In this Class? (WIHIC), to assess students' perceptions of the learning environment; and a scale to assess students' enjoyment of their English classes (both of which are described in Chapter 4). To examine

differences in pre–post scores, MANOVA with repeated-measures were performed using the sample of 2,417 students. The results indicated that there were statistically significant (p<0.01) differences for six of the seven WIHIC scales (namely Student Cohesiveness, Teacher Support, Involvement, Finding Reference, Task Orientation and Cooperation) and the attitude scale before and after the teacher professional development programme. Despite the statistical significance, the effect sizes were small (according to Cohen's criteria), suggesting limited educational significance.

Given that the observations made at Phase 3, which indicated that there may be differences in the extent to which the teachers in urban and rural classrooms translated the ideas from the MGMP Empowerment programme into practice (reported in Section 5.3), it was important to examine whether the MGMP Empowerment programme was differentially effective for teachers in these two localities. A two-way MANOVA, involving the seven learning environment scales and one attitude scale as the set of dependent variables and the two independent variables involved a two-level variable based on location (urban and rural) and the repeated measure (pre-test and post-test).

The findings indicated that for six of the seven learning environment scales (the exception being the Involvement scale) and the attitude scale, there existed a statistically significant interaction between testing occasion and location. For the learning environment scales and the attitude scale, there existed a locality gap at the pre-test which grew at the post-test. For all learning environment scales and the attitude scale with statistically significant interaction for both the pre-test and post-test, students in urban areas scored higher than their rural counterparts. In all cases, the magnitude of the effect sizes indicated that students in urban schools (effect sizes ranging from 0.06 to 0.21) were higher than their counterparts. In addition, for students in urban schools, all pre–post changes were greater, and positive in direction, whereas for students in rural schools, the pre–post changes were small, and in some cases, were negative in direction. These pre–post changes suggest that the programme was differentially effective for schools in rural

and urban areas. An examination of the contextual factors helped to understand the possible reasons for these disparities.

The results of this study indicated that contextual factors, which might promote or inhibit the teachers' efforts to implement the ideas from the teacher professional development programme, might exist (Phase 5). Three overarching factors appeared to have influenced the teachers' efforts to translate the ideas of the professional development programme in to practice: the instructional context; insufficient support provided by the school administrators; and the nation-wide leaving examination. The two case study teachers in urban schools that made changes to their practice felt that the status of the school and the socio-economic status of the students encouraged them to try out the new ideas. In addition, the students' good behaviour and their motivation to study made it easier for the teachers to try out the suggestions imparted during the professional development. In contrast, the teachers from rural schools claimed that issues related to student discipline, large classroom sizes and inadequate resources and facilities hindered their attempts to implement the new ideas in their classrooms.

All six case study teachers indicated that there was inadequate practical support from either the principals or superintendents. This lack of support was found to hamper the teachers' efforts and their willingness to try out new ideas. In all cases, the teachers felt that the principals did not create a culture of change, there was no support as they tried to implement the changes and that the opportunities for teachers to work collaboratively to help each other to plan for change were not available. Finally, the teachers felt that the constructive feedback from principal and superintendents were seen to be hindrances to their attempts to make changes to their teaching practice.

The six case study teachers all viewed the nation-wide leaving examination as a major hindrance to adopting more the student-centred approaches. Although student-centred approaches were mandated in the national curriculum, it appears that the high-stakes national examination forced the teachers to continue to use

teacher-centred approaches. These approaches were viewed by the teachers as being more effective in preparing students for the final examination.

Chapter 6 presents the discussion and conclusion.

Chapter 6

DISCUSSION AND CONCLUSION

6.1 Introduction

The overarching aim of my study was to evaluate the effectiveness of the MGMP Empowerment programme. The programme was intended to provide the opportunity and support to advance the practices of the participating teachers. It was expected that, by participating in the programme, teachers would improve their teaching practices in ways that were more student-centred in order to meet the standards of the new national curriculum.

This chapter concludes the thesis using the following subheadings:

- Discussion of Findings (Section 6.2);
- Contributions of the Study (Section 6.3);
- Limitations of the Study (Section 6.4);
- Future Research (Section 6.5); and
- Final Comment (Section 6.6).

6.2 Discussion of Findings

The study evaluated the effectiveness of the MGMP Empowerment Programme, in terms of: teachers' views towards the delivery and utility of the programme; changes in the teachers' teaching strategies; and changes in students' learning experiences and attitudes towards English. To accomplish this evaluation, five research objectives were proposed, the findings for which are summarised and discussed below. Section 6.2.1 summarises the development of a model suitable for evaluating teacher professional development in the Indonesian context. Section

6.3.2 summarises the findings for the reliability and validity of instruments used as part of the evaluation. Section 6.3.3 summarises the effectiveness of the programme in terms of changes in the teaching strategies that the teachers used and changes in students' learning experiences and attitudes towards English. Given that the data collected to examine the effectiveness of the programme indicated that the programme could have been more effective for teachers in urban than their counterparts rural schools, Section 6.3.4 summarises the findings for the analyses of data to examine the differential effectiveness of the professional development programme. Finally, Section 6.3.5 summarises the contextual factors that could have promoted or impeded the translation of the professional development programme into practice.

6.2.1 The Development of a Model Suitable for Evaluating Teacher Professional Development in the Indonesian Context

To enable me to effectively evaluate the teacher professional development programme, it was important to develop a model of evaluation that was contextually relevant and took into consideration the views of students, as the major stake holders in the education process. Therefore, the first research objective sought to develop a suitable model. My model drew on three existing models that were originally developed by Mathison (1992), Guskey (2000) and Fishman, Mark, Best and Tal (2003). An important feature of my model was the inclusion of students' perceptions of their learning experiences and their attitudes to provide a more holistic evaluation that took into consideration whether these changed over the course of the one year programme.

This new model involved five phases. The first phase involved the collection of important baseline data which examined the teaching strategies used by the teachers and the students' attitudes and perceptions of their learning experiences prior to the commencement of the professional development programme. The second phase involved evaluating teachers' attitudes towards the delivery and utility of the professional development programme and whether they felt that the

programme had provided them with the knowledge and skills required to change their teaching practices in ways that were more student-centred. The third phase involved examining the extent to which the ideas and concepts, imparted during the professional development programme, were translated into practice by the teachers. The fourth phase examined whether (at the end of the programme) students perceived their classroom learning experiences differently to the way they did before the programme commenced and whether their attitudes toward their English classes had changed. The fifth and final phase of the model involved examining contextual factors that might promote or impede the translation of ideas imparted during the professional development programme.

This model provided a framework that guided the collection of data that was used to evaluate the effectiveness of the professional development programme.

6.2.2 Reliability and Validity of the Surveys

Part of the evaluation of the professional development programme involved the collection of data using two surveys (one to assess students' perceptions of the learning environment and one to assess students' attitudes towards their English classes). Therefore, my second research objective involved modifying, translating and validating the two surveys to ensure that the results could be used with confidence.

The two instruments, the What Is Happening In this Class? (WIHIC) (used to assess the students' perceptions of their learning environment) and the Enjoyment of English Class scale (to assess the students' attitudes towards the subject), were administered at Phase 1 and again at Phase 4 of the evaluation. To ensure their suitability for use in lower secondary classes in Indonesia, both instruments were modified and translated into Indonesian. First, the scales of the surveys were inspected to ensure their relevance to English language classes. During this process, one of the original scales of the WIHIC (Investigation) was removed and, for the purpose of this study, the Finding Reference scale was developed. Second,

individual items were scrutinised to ensure that the language was suitable and the items were relevant to English language classes. Once the surveys had been modified, they were translated into the Indonesian language using the rigorous process of back translation (Bracken & Barona, 1991; Brislin, 1970, 1980; Chapman & Carter, 1979). The following sections overview the validity and reliability of the WIHIC questionnaire (Section 6.2.2.1) and attitude scale (Section 6.2.2.2).

6.2.2.1 Validity and Reliability of the Indonesian Version of WIHIC

Data collected from 2,417 students in 66 classes were analysed in various ways to investigate the validity and reliability of the Indonesian version of WIHIC. As a first step, principal axis factor analysis with oblique rotation was used to examine the factor structure of the modified WIHIC separately for the pre-test and post-test data. Of the 56 items, eight were found to be problematic and, to strengthen the factor structure, were omitted and not included in any further analysis. The factor analysis resulted in a refined version of the WIHIC, comprising of 48 items in the seven *a priori* scales. Each of the 48 items had factor loadings of at least 0.30, with the exceptions of two of items on the pre-test, these being, Item 7 for the Student Cohesiveness scale (which loaded on the Cooperation scale) and Item 49 for the Equity scale (which loaded on the Student Cohesiveness scale).

Estimates of the internal consistency of each WIHIC scale were determined using Cronbach's alpha coefficient. The Cronbach alpha coefficients using the individual student as the unit of analysis, ranged from 0.69 to 0.89 for the pre-test and from 0.73 to 0.91 for the post-test. Using the class mean as the unit of analysis, the alpha reliability coefficients ranged from 0.91 to 0.98 for the pre-test and from 0.92 to 0.98 for the post-test.

The discriminant validity was calculated to determine the extent to which the scales were independent of each other. The results of the component correlation matrix, obtained from oblique rotation, indicated that the highest correlation between any of the WIHIC scales was 0.46, which met the requirements of discriminant validity according to Brown (2006). The discriminant validity for the pre-test and the post-

test data support that the scales of the Indonesian version of the WIHIC measured distinct constructs for both administrations.

To examine whether the scales of the Indonesian version of the WIHIC were able to differentiate between classes, one-way ANOVA with classroom membership as the independent variable was used. Each of the seven scales of the WIHIC, for both the pre-test and the post-test administrations, were able to differentiate significantly (p<0.01) between classes. The values for the eta^2 statistics (the proportion of variance) ranged from 0.19 to 0.25 for the pre-test and from 0.21 to 0.30 for the post-test. Overall, the ANOVA results provided evidence that the scales of the Indonesian version of the modified WIHIC were able to differentiate between different classes.

The reliability and validity of the Indonesian version of the WIHIC replicated numerous past studies that used the WIHIC in western countries (Aldridge & Fraser, 2000; Allen & Fraser, 2007; Biggs, 2008; Dorman, 2003; Ogbeuhi & Fraser, 2007; Velayutham et al., 2011; Wolf & Fraser, 2008) and non-western countries (Aldridge, Fraser & Huang, 1999; Aldridge, Fraser & Ntuli, 2009; Chionh & Fraser, 2009; Khoo & Fraser, 2008; Kim, Fisher & Fraser, 2010; Koul & Fisher, 2005). In particular, these results are consistent with results obtained in the past studies that used the WIHIC questionnaires in Indonesia (Fraser, Aldridge & Adolphe, 2010; Margianti, Aldridge & Fraser, 2002; Wahyudi & Treagust, 2004).

6.2.2.2 Validation of the Attitude Scale

To examine the reliability of the attitude scale, the internal consistency reliability was calculated. With the student as the unit of analysis, the alpha coefficient was 0.90 for the pre-test and 0.91 for the post-test. With the class mean as the unit of analysis, the alpha coefficient was 0.97 for the pre-test and 0.98 for the post-test. These high coefficients supported the reliability of the Indonesian version of the attitude scale. These results were consistent with the results of past studies that have included the enjoyment scale in Indonesia (Aldridge, Fraser & Adolphe, 2010; Margianti, Aldridge & Fraser, 2004). These results suggested that both the

Indonesian version of the WIHIC and attitude scale were reliable when used in English classes in Indonesia. It was considered, therefore, that the findings of subsequent analysis could be used with confidence.

6.2.3 Effectiveness of the MGMP Empowerment Programme

Using the new evaluation model as a framework for the collection and analysis of data, the third research objective sought to evaluate the professional development programme in terms of: teachers' views of the programme and their views about whether the programme had improved their knowledge and skills with respect to teaching (Phase 2 of the evaluation model); whether there were changes in the teaching strategies that the teachers used (Phase 3 of the evaluation model); and whether students' perceptions of their learning environment and attitudes towards English changed over the course of the one year programme (Phase 4 of the evaluation model).

6.2.3.1 Phase 2 of the Evaluation

To examine teachers' views of the programme and whether they felt that the programme had improved their knowledge and skills with respect to implementing the new teaching strategies (Phase 2 of the evaluation model), data were gathered from teacher reflective journals and in-depth interviews with teachers (n=33). The results indicated that all of the teachers were positive about the facilities that were provided during the programme.

In terms of the delivery of the programme, however, the teachers felt that there were problems with respect to: the calibre of two of the instructors; and the time required to complete the programme. Interviews with the teachers, suggested that most of them felt that two of the four trainers were not adequately prepared. In particular, the teachers felt that the presentations made by two of these instructors failed to provide them with clear teaching suggestions and did not explain the teaching materials clearly. The teachers further agreed that, most of the time, these two instructors used one-way lectures rather than workshops; leaving them with

the impression that the trainers had inadequate knowledge of both the curriculum and of appropriate teaching skills. These two trainers were both lecturers, drawn from local universities, neither of whom had taught in the secondary school system. As a result, these university lecturers were not familiar with the problems faced by teachers in their daily teaching practise.

According to Holloway (2000), professional development is unlikely to be successful if the presenters are not fully conversant with the relevant practices and materials that they are exposing the teachers to. In response to this problem, it is recommended, therefore, that future planning for teacher professional development pay particular attention to involving experienced teachers who are not only qualified in terms of knowledge of the subject matter, but also have extensive classroom experience. It is recognised that many professional development instructors are not currently experienced teachers so, it is recommended that the teacher professional development provider, together with local governments, provide incentives for experienced teachers to become instructors.

The teachers also complained that the time available to cover the programme materials was too tight. As a result, the teachers suffered from fatigue and the sessions generally felt rushed. The teachers felt that they were overloaded with information at the end of each day. In addition, the teachers felt that this rushed timetable left them with insufficient opportunities to hold discussions with colleagues. As Guskey (2000) pointed out, one of the most crucial aspects of planning teacher professional development is the provision of time for discussion with colleagues. These sessions should include discussions about ways to improve their teaching strategies, difficulties that they are experiencing and developing solutions for those difficulties, planning together and looking for ways to collaborate. In response to this problem, when planning future teacher professional development, it is recommended that, to enhance the outcomes of the programme, consideration given to the provision of adequate discussion time and the

incorporation into the programme of various activities to allow teachers to have meaningful discussions.

Despite their less than positive views towards the delivery of the programme, the case study teachers indicated that, generally, the professional development programme had expanded their knowledge and skills. The teachers expressed that they were keen to improve their teaching and were enthusiastic to incorporate the ideas, imparted during the professional development programme, into their classroom practice. Finally, the results of the evaluation conducted in Phase 2 revealed that the teachers generally felt that, despite its limitations, the professional development programme was valuable in terms of meeting colleagues and sharing ideas and experiences with others.

6.2.3.2 Phase 3 of the Evaluation

Phases 1 and 3 of the evaluation model involved the collection of interview and observation data to examine the extent to which the ideas imparted during the professional development programme were translated into classroom practice. Video analysis of four hours of observations in each of the classrooms of the six case study teachers suggested that only two of the six case study teachers, both of whom taught in urban schools, changed their pedagogies in ways that were aligned with the ideas imparted during the professional development programme. These teachers both included collaborative group work in their teaching which they had not used previously.

Despite their enthusiasm to implement the new ideas into their classroom practice (Phase 2 of the evaluation), none of the other four teachers showed any signs of change from their teacher-centred practice. One of these four teachers was teaching in an urban school and, according to his interviews, he viewed the new teaching methods (proposed during the professional development programme) as ineffective. In his opinion, these methods were not helpful in preparing his students for the nationwide examination and despite his original intentions (expressed at Phase 2) he was not prepared to change his teaching strategies.

Observations in the classes of the three teachers who taught in rural schools suggested that the teaching strategies remained unchanged. These teachers continued to use a teaching format that consisted of one-way lectures followed by a question and answer session, during which the teachers asked the questions. During the lessons, the teachers generally explained concepts while the students copied notes from the board. These sessions were generally followed by a written exercise. Even though there was one occasion during which the teacher asked the students to work in pairs, this attempt was quickly abandoned when the students became unruly.

The results of the present study are similar to previous studies that evaluated teacher professional programmes in Indonesia. A review of the literature revealed that past teacher professional development programmes with a variety of delivery approaches have been conducted in Indonesia since 1970s, but there has been little evidence to demonstrate the effectiveness of the programmes (Fuller, 1987). The results of these studies suggested that the teachers who attended the professional development activities generally were unable to implement the new ideas in their classrooms (Fuller, 1987).

In my study, the ideas that the teachers were exposed to during the professional development programme were viewed, by some of the teachers, as inappropriate for their classroom contexts and, as a result, they found it difficult to incorporate these ideas into their day-to-day teaching. This replicates another study in which surveys were administered in 10 provinces in Indonesia, the results of which indicated that teachers persisted in their domination of the classroom with little attention being paid to the students' learning experiences (Mahady, Wardani, Irianto, Somerset & Nielsen, 1996). Also Naim's (1995) study in Indonesia found that the teachers retained a teacher-centred approach even after their completion of a teacher training aimed at assisting them to move to a student-centred approach. In another study conducted in Indonesia, Thair and Treagust (2003) reported that, while successful in establishing network of teacher trainers and a standardised

approach to teacher professional development, classroom observations indicated that the teaching practices of the teachers remained unaffected.

6.2.3.3 Phase 4 of the Evaluation

Phases 1 and 4 of the evaluation model involved examining changes in students' perceptions of the learning environment and attitudes. To do this, the Indonesian version of the WIHIC and attitude scales were administered to 2,417 students at the beginning of the programme and again (one year later) at the end of the programme.

A one-way MANOVA for repeated measures (with the WIHIC and attitude scales serving as dependent variables and the pre-test–post-test as the repeated measure) was conducted. In addition, the effect size was calculated, to assist in determining whether a difference might be educationally significant. Using Wilk's lambda criterion, differences for the set of WIHIC and attitude scales were found to be statistically significant and, therefore, the univariate one-way ANOVA was interpreted for each WIHIC and attitude scale. The results indicated that there was a statistically significant (p<0.01) pre–post differences for six of the seven WIHIC scales (namely, Student Cohesiveness, Teacher Support, Involvement, Finding Reference, Task Orientation and Cooperation) but not for the attitude scale. All scales showed an improvement between the pre-test and the post-test. Despite there being statistically significant differences for those six scales, the effect sizes were too small to be of educational significance, according to Cohen's (1992) criteria.

According to Kilpatrick (2000) teachers can be resistant to changing their instructional practices. Teachers who claim they have totally revamped their instruction may, when observed, appear to be teaching in rather traditional fashion Teacher's commitment to change must be taken into consideration when developing professional development. Johnson (2001) argues that effective elements of teacher professional development include incentives such as recognition and additional certification that might be used to motivate teachers. As the MGMP Empowerment

programme was not incorporated into the mainstream of educational policy, teachers' participation and efforts in classroom improvement were not recognised. This lack of recognition may have inhibited the teachers from more fully committing to the suggested changes. This problem suggests the need to integrate teacher professional development with career development in order to be sustainable. The provision of recognition through certification and financial incentives given to the teachers upon their completion of the programme are examples of how this might be achieved through institutional support.

6.2.4 Differential Effectiveness of the Teacher Professional Development for Teachers in Urban and Rural Schools

During Phase 3 of the evaluation, the analysis of the data suggested that differences, in terms of the teachers' use of new knowledge and skills, might exist between teachers in urban and rural schools. Therefore, the fourth research objective sought to investigate whether the MGMP Empowerment programme was differentially effective for teachers in urban and rural schools in terms of: a) students' perceptions of their learning experiences; and b) students' attitude towards English classes.

To investigate whether the professional development programme was differentially effective for teachers in urban and rural schools, a two-way MANOVA was used. For the two-way MANOVA, the independent variables were the testing occasion (pre-test and post-test repeated measure) and the location (rural and urban), and the dependent variables were the WIHIC and attitude scales. To examine the magnitude of the differences, the effect size was calculated separately for urban and rural classes for each WIHIC and attitude scale. In addition, qualitative data were used to provide a better understanding of the quantitative data.

Examination of the average item means clarified that, for the WIHIC scales, the students in urban areas had more positive views than their rural counterparts for both the pre-test and post-test. The two-way ANOVAs for testing occasion, with

control for location indicate statistically significant pre–post improvements for six of the seven learning environment scale but not the attitude scale. In all cases, the effect sizes (calculated using the eta^2) were small. The ANOVA results for whether differences exist for students in rural and urban locations, regardless of testing occasion, indicated statistically significant differences for all seven learning environment scales and attitude scale. In all cases, the directions were positive and the proportion of variance was small.

To examine the differential effectiveness of the MGMP Empowerment programme, the interactions between the testing occasion and location, the two-way ANOVAs were inspected. The results showed that a statistically significant interaction existed for six of the seven learning environment scales and the attitude scale – indicating that the independent interactions of the pre–post differences were valid only for the Involvement scale.

For the six WIHIC scales with positive interactions, the results indicated that, although a locality gap existed for the pre-test, this gap was larger for the post-test. In all cases, the pre–post improvements for students in urban schools were positive and greater than for students in rural schools. In addition, the effect sizes for the pre–post differences were larger for students in urban schools (ranging from 0.06 to 0.21 standard deviations) than for students in rural students (ranging from 0.00 to 0.06 standard deviations).

Qualitative information, collected using in-depth interviews with case study teachers and their students, supported the quantitative findings and provided insights into the differences. The results suggested that, for four of the learning environment scales (Student Cohesiveness, Teacher Support, Task Orientation and Cooperation) the use of group work in urban schools may have contributed to the increased gap in the post-test scores. It would appear that the disparity between the use of Finding Reference scale for students in urban and rural schools may be due to the lack of resources available to students in rural schools. Because these resources were available to students in urban (both at home and in class), the

teachers was more likely to use them than their counterparts in rural schools whose students did not have access to these resources. In terms of the locality gap for the Equity scale, interviews with teachers and students suggested that, teachers in rural schools are more likely to expend their energy teaching students who were considered to be capable or motivated (often ignoring the needs of the other students), whereas the teachers in urban schools were more inclined to tend to the needs all of the students in the class.

The results also suggested that there was a statistically significant interaction for the Enjoyment of English Class scale. The results indicated that, in urban schools, students enjoyment of English classes improved (effect size = 0.11 standard deviations), whereas the enjoyment of English classes for their rural counterparts declined marginally (effect size= 0.06 standard deviations).

According to interviews with students and teachers, differences in Enjoyment of English Classes for urban and rural students was influenced by a number of factors including the teaching strategies used by the teachers, the materials that were presented during English classes and the value that the students held for learning English. It would appear that the students in urban schools enjoyed their English classes more than students in rural classes because their teachers were friendlier and they viewed learning English as important to their future. As a result these students tended to be more engaged than their counterparts in rural schools.

The results of my study suggest that there were disparities between the usefulness of the knowledge and skills imparted during the programme for teachers in urban and rural schools. These results replicated Wahyudi's and Treagust's (2004) study which found differential effectiveness of the implementation of a new science curriculum by exemplary teachers in urban and rural schools in Indonesian lower secondary schools. In this study, it was reported that after completing a teacher professional development programme, the science teaching practices in urban schools changed to be more student-centred. Whereas the teaching practices in

rural schools continued to be teacher-centred, largely due to the limited facilities and resources available in the rural schools (Wahyudi & Treagust, 2004).

My results indicate that the teachers in rural schools were disadvantaged by conditions such as large class sizes, student discipline problems and inadequate resources and facilities, all of which made it difficult to implement the student-centred strategies. Guskey and Yoon (2009) suggest that teacher professional development must provide teachers with strategies that are able to be adapted to different classroom contexts. The results imply that the strategic interventions and materials, suggested during the MGMP Empowerment programme, would need to be modified to suit specific local contexts if they were to be implemented successfully by teachers in rural schools. This could require an alternative professional development approach that tailors the professional development programme to meet the needs of individual schools. In this case, teacher professional development would need to be carefully planned to equip the participants from urban and rural schools with skills that enable them to use the materials within the contexts of their schools.

6.2.5 Contextual Factors Promoting or Impeding Implementation at the Classroom Level

The fifth research objective was delineated to examine whether contextual factors exist that might promote or impede changes in teaching practice. To examine the contextual factors that influenced the implementation of the MGMP Empowerment programme (Phase 5 of the evaluation model), in-depth interviews with the six case study teachers were used. The results suggested that these teachers were generally positive about the professional development programme and were keen to find ways to improve their teaching. All of six of the case study teachers felt that the programme had increased their knowledge and skills and that they were excited at the prospect of trying out the ideas (Phase 2 of the evaluation). However, analysis of qualitative information (collected using in-depth interviews with the six case study teachers) indicated that there were three overarching factors that influenced

the likelihood of teachers translating the ideas, promoted during the professional development programme, into practice: the instructional context; organisational support; and the existing nation-wide examination.

The instructional context, including class size, availability of resources and the socioeconomic status and the behaviour of the students, appears to have influenced the extent to which the teachers were either able or willing to implement the new ideas. The two case study teachers who were teaching in urban schools and who made changes to their teaching practices, indicated that their willingness to implement the new ideas was due, in part, to their students' good behaviour and the social and economic status of the families. The teachers also made clear that the students' attendance at after school tuition classes made trying out the new ideas possible. The teachers felt that, because students attended after-school tuition, they were not reliant on the teacher providing them with the content needed to pass the examinations and could, therefore, try the new ideas without fear of their students failing.

In contrast, all three of the case study teachers who were teaching in rural schools were reluctant to use the new approaches. Interviews with these teachers indicated that they all believed that the traditional teacher-centred approaches better suited their students than student-centred approaches. It appears that, despite their best intentions, the large class sizes, student discipline and classroom management issues and the lack of resources made it difficult for them to implement changes to their practise.

The second factor that appeared to influence implementation was the support, or lack thereof, from the organisation. The MGMP Empowerment programme recommended that teachers work together to plan activities and lessons, share resources and accompany each other to their classrooms. However, in all cases, the organisation did not assist with timetabling to facilitate this. In addition, the principals in rural schools were either unwilling or unable to assist teachers with their requests for much needed resources that may have been beneficial to the

implementation of ideas. These teachers went further to suggest that there was no attempt by the organisation to create an environment in which teachers felt supported and encouraged. In all cases, the teachers felt that the principal and superintendent were unaware of what they were doing and had no knowledge of the requirements made of them during the programme. Finally, the teachers were all cognisant of the fact that neither the principal nor superintendent was available for personal support.

Research indicates that the ability to implement skills learned outside the classroom into classroom practice needs practical support to sustain it (Kilpatrick, 2000). The results of this study suggest that the lack of support from the principals and higher level administrators was one of the constraints placed on the teachers as they implemented the suggestions of the MGMP Empowerment programme. Therefore, it is important that, to improve the likelihood of translating skills and knowledge into practices, the MGMP Empowerment Programme pays attention to establishing collaboration between teachers, head-teachers, principals and education department personnel that positively reinforce the intended outcomes of the professional development strategy.

The third factor that appeared to be a major impediment to the teachers' willingness to implement changes was the nation-wide leaving examination. The teachers unanimously agreed that the students' scores on this high-stakes examination were paramount – not only for the students but also for their own reputation and career. The teachers were, without exception, uncertain about whether that the student-centred approach would be beneficial in preparing their students for the examination. This, coupled with their fear of failing to cover the syllabus because of the tight time-schedule, influenced the teachers' decisions to continue with teacher-centred approaches rather than implementing the student-centred approaches mandated by the new curriculum and suggested during the professional development programme.

The results of my study reflect past studies that have illustrated that, if there is tension between policy rhetoric and the reality that teachers face in their day-to-day work, then the implementation of national curriculum innovation mandates are likely to be impeded (Fullan, 2007). In my study, a major constraint to translating the professional development ideas into practice was the need to prepare students for the high-stakes nation-wide examination. These findings imply that, when planning teacher professional development, the broad framework of the national educational policy needs to be considered. If, as mandated in the national curriculum, teachers are to use more student-centred methods then the Indonesian government will need to consider transferring greater management and decision making authority to the provincial and regional levels to allow them to design curriculum and evaluation procedures that meet the local needs.

6.3 Significance of the Study

The study has made a distinctive theoretical contribution with the development of a new model for evaluating professional development that is suited to the Indonesian context. The model is consistent with research related to professional development and takes into account a range of factors that influences the success of a professional development programme at various stages of the implementation

This study has made a significant contribution to the field of learning environments by demonstrating the value of including learning environment assessment as an integral part of an evaluation of the effectiveness of a teacher professional development programme. The use of the Indonesian version of the learning environment questionnaires has helped to determine whether students' learning experiences have changed as the impact of the teacher professional development programme. In this sense, my study is consistent with past research (Martin-Dunlop & Fraser, 2008; Pickett & Fraser, 2009).

My study has established the robustness and applicability of learning environment instruments that were used as an integral part of the new evaluation model of

teacher professional development. These instruments were modified to ensure their suitability for use in English classes in Indonesia and can be used in future evaluations of professional development This contribution is significant within the field of learning environments as it extends the range of learning environment instruments and the research that has been conducted within the Indonesian context.

The results of the present study have provided practical implications for a variety of stakeholders including policy-makers, teacher professional development providers and teachers. The students' perceptions of classroom learning can provide critical feedback for evaluating the effectiveness of the teacher professional development.

The results of this study provide important information to the professional development providers who were central in this evaluation. The results could provide useful feedback (in order to understand the strengths and weaknesses) that can be used to improve the design of future professional development programmes in Indonesia. In particular, the emphasis of future improvements would be on improving the effectiveness of teacher professional development outcomes for teachers in rural schools.

6.4 Limitations

The present study involved students' perceptions of the learning environment and their attitudes towards English as part of the measures to evaluate the effectiveness of a teacher professional development programme for lower secondary English teachers. In interpreting the findings, there are a number of limitations that should be considered. First, in this study, as the researcher, I was conscious that the data gathered from a relatively small number of case study teachers (n=6) needed to be treated with caution. Therefore, the generalisability of findings from this study is limited. It is recommended, therefore, that future research involve a larger number of case study teachers.

However, a larger number of teachers (*n*=33) were interviewed to overcome this problem. In addition, a large quantitative student sample (*N*=2,417) was involved to help to improve the generalisability of the findings.

Second, limitations arose as a result of time constraints. The administrations of the pre-test and the post-test of learning environment and attitudes surveys involved a relatively short timeframe (one academic year). Past research related to the evaluation of teacher professional development has found that, to better gauge the success of an innovation, it was desirable to re-administer two years after the conclusion of the programme (Fullan & Stiegelbauer, 1991; Guskey, 2002). Given that the first six months of programme implementation is likely to be characterised by high anxiety and confusion among teachers (Huberman's, 1981), this one-year timeframe for the pre–post administration could have contributed to the limited changes in teaching practice.

The third limitation was the absence of data related to student cognitive outcomes. Literature related to teacher professional development emphasises that, to examine the impact of effective teacher professional development programme, improvements in student achievement should be considered (Joyce & Showers, 2002; Fullan et al., 2006). Difficulty in accessing data related to the students' achievement scores made it impossible to address this aspect. Although it is acknowledged that data related to students' affective outcomes are not considered to be as powerful, data related to students' attitudes was used to help to gauge the implementation of the professional development programme.

As described in Section 5.2, in this study I interviewed teachers to examine their views on the professional development experiences. One possible concern is that, although most teachers attend teacher professional development programme because they want to become better teachers, some teachers might attend this programme to fulfil government mandates to retain their employment. Therefore, this limitation should be considered when interpreting the results.

In explaining the differential effectiveness of the MGMP empowerment programme for teachers in urban and rural schools, consideration was not given to the differences in exposure to native English speakers in both settings. As it is possible that this may have impacted on students' motivation to learn English, this limitation should be taken into consideration when interpreting the results.

Time restrictions precluded the researcher from conducting sustained observations in the classes of the case study teachers. Given that teacher behaviour may have differed over a longer period of observations, this limitation should be considered when interpreting the results. Clearly, for future research of this nature, it would be desirable to include a richer range of sustained classroom observations.

The absence of a control group could be considered to be a limitation. By including an investigation of lower secondary English classrooms that were not participating in the MGMP Empowerment programme might have been useful but, given the time constraints of my study and the large amounts of data collected from participating teachers, this was not possible.

6.5 Future Research

Future research involving the evaluation of teacher professional development has considerable potential; and the use of learning environment instruments could play an important role in that growth. Future studies could use the newly-modified WIHIC and attitude scales for the evaluation of teacher professional development programmes. In particular, the continuing emphasis on enhancing English instruction in Indonesia provides scope for the use of these instruments as a component of future evaluations.

Future studies could involve replicating my study in other parts of Indonesia in order to increase the validity and generalisability of the findings. Using the Indonesian version of the WIHIC and attitude scales with participants of the MGMP Empowerment programme in other districts could add confidence in my results.

Future studies could include the evaluation of a professional development programme with a comparison group of teachers who do not participate in the professional development. The addition of a control group to the study could shed light on the sensitivity of the learning environment and attitude instruments to changes in the students' perceptions of their experiences at the classroom level.

Future studies might incorporate a longitudinal design, with the inclusion of a non-participant comparison group. Increasingly teacher professional development is embracing a sustained approach, thus investigating the learning environment in classrooms over time might provide useful information in evaluating the impact of professional development efforts. Coupled with this longitudinal design could be the inclusion of measures of student achievement in English, thus allowing a more robust attempt to use student outcomes for professional development evaluation.

Another potential area of study could involve incorporating the underlying constructs of learning environment instrument into the framework of the professional development programme design (i.e. using scale descriptors from the instruments when establishing the goals of a professional development programme). The study could then examine the effect of an overt effort to clarify the characteristics of a learning environment on the changes actually produced in the classroom over time.

The potential for future educational research that includes the WIHIC and attitude scale is considerable. In addition to adding to academic knowledge about dynamic classroom systems, learning environment instruments could be used by individual teachers to enhance their classroom performance and, hence, the cognitive and affective performance of their students.

6.6 Concluding Remarks

My evaluation of the MGMP Empowerment programme revealed that the programme was more effective for teachers in urban schools than their counterparts in rural schools. The results indicated that the MGMP Empowerment programme was not able to cater to the needs of teachers in rural schools who had to deal with issues such as large classroom size and lack of resources and facilities.

Effective professional learning requires sufficient time and resources, as well as supportive structures from colleagues and higher level administrators (Guskey, 2002). In my study, the results indicate that these elements, required for the success of the MGMP Empowerment programme, may have been lacking. To better support teacher change, my results imply that the effectiveness of the MGMP Empowerment programme would be enhanced if good supporting materials and practical supports, such as providing a culture of change, opportunities for teachers to collaborate and constructive feedback, were in place during the implementation.

The results of my study highlight the need to view teacher professional development as an ongoing activity. As such, there needs to be a sustained timeframe that is built into the ongoing work of teaching which needs to be sensitive to contextual conditions. Careful attention to follow up activities, after the completion of teacher professional development, is important. Collaborative support from colleagues and problem solving with respect to student learning and the school context must be carefully considered. Lastly, while providing effective teacher professional development is challenging, it must be seen as being at the core of teachers' professional work as they develop expertise in their practice.

References

Abell, S. K., Lannin, J. K., Marra, R. M., Ehlert, M. W., Cole, J. S., Lee, M. H., et al. (2007). Multi-site evaluation of science and mathematics teacher professional development programs: The project profile approach. *Studies in Educational Evaluation, 33*, 135-158.

Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research, 3*, 101-134.

Aldridge, J. M., & Fraser, B. J. (2008). *Outcomes-focussed learning environments: Determinants and effects*. Rotterdam, the Netherlands: Sense Publisher.

Aldridge, J. M., Fraser, B. J., & Huang, I. T. C (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *The Journal of Educational Research, 93*, 48-62.

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

Aldridge, J. M., Fraser, B. J., Bell, L. M., & Dorman, J. P. (2012). Using a new learning environment questionnaire for reflection in teacher action research. *Journal of Science Teacher Education, 23*, 259-290.

Aldridge, J. M., Fraser, B. J., & Sebela, M. P. (2006). Using teacher action research to promote constructivist learning environments in South Africa. *South African Journal of Education, 24*, 245-253.

Aldridge, J. M., Fraser, B. J., Taylor, P. C., & Chen, C. C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22*, 37-55.

Afari, E., Aldridge, J. M., Fraser, B. J., & Khine, M. S. (in press). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research*.

Allen, D., & Fraser, B. J. (2007). Parent and student perceptions of classroom learning environment and its association with student outcomes. *Learning Environments Research, 10*, 67-82.

Asghar, M., & Fraser, B. J. (1995). Classroom environment and attitudes to science in Brunei Darussalam. *Journal of Science and Mathematics in S.E. Asia, XVIII*, 41-47.

Badan Pusat Statistik. (2010). *Statistics Indonesia*. Retrieved 30 October, 2011, from www.bps.go.id

Baird, J. R. (1992). Collaborative reflection, systematic enquiry, better teaching. In T. Russel & H. Munby (Eds.), *Teachers and teaching: From classroom to reflection* (pp.33-48). New York: Palmer Press.

Balitbangdiknas. (2009). *Laporan Hasil Ujian Nasional*. Retrieved 23 May, 2009, from http://puspendik.info/v4/index

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-31). San Francisco, CA: Jossey-Bass.

Bagozzi, R., & Burnkrant, R. (1979). Attitude organization and the attitude-behavior relationship. *Journal of Personality and Social Psychology, 37*, 913-929.

Biggs, E. M., & Fraser, B. J. (2008). Evaluation of a professional development program on integrating technology into middle schools: Classroom environment and student attitudes. Unpublished Doctoral Thesis, Curtin University, Perth, WA.

Bracken, B. A., & Barona, A. (1991). State of art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International, 12*, 119-132.

Brislin, R. (1970). Back translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*, 185-216.

Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology*. *(pp.* 389-444*)* Boston: Allyn and Bacon.

Broad, K., & Evans, M. (2006). A review of literature on professional development content and delivery modes for experienced teachers: Report for the Ontario Ministry of Education. Ontario, Canada: Ministry of Education.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guildford Press.

Brown-Easton, L. (2004). *Powerful design for professional learning*. Oxford, Ohio: NSDC.

Chapman, D. W., & Carter, J. F. (1979). Translation procedures for the cross-cultural use of measurement instruments. *Education Evaluation and Public Analysis, 1*, 71-76.

Chionh, Y. H., & Fraser, B. J. (2009, February). Classroom environment, achievement, attitudes and self-esteem in geography and mathematics in Singapore. *International Research in Geographical and Environmental Education, 18*, 29-44.

Cho, J., Yager, R. E., Park, D. Y., & Seo, H. (2004). Changes in high school teachers' constructivist philosophies. *School Science and Mathematics, 90*, 400-406.

Clarke, D. M., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*, 947-967.

Clayton, A. M., & Thorne, T. (2000). Diary data enhancing rigour: Analysis framework and verification tool. *Journal of Advanced Nursing, 32*, 1514-1521.

Coakes, S., & Steed, L. (2007). *SPSS version 14.0 for windows: Analysis without anguish.* Milton, Australia: John Wiley & Sons.

Cochran-Smith, M., & Lytle, S. (1993). Inside/Outside: Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education, 24*, 249-306.

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*, 155-159.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. New York: Routledge.

Corcoran, T. B. (1995, June). *Helping teachers teach well: Transforming professional development (Policy Brief No. RB-16).* New Brunswick, NJ: Consortium for Policy Research in Education.

Cotterall, S. (1995). Readiness for autonomy: Investigating learner beliefs. *Systems, 32*, 195-205.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*, 1-9.

Crano, W., & Prislin, R. (2006). Attitudes and persuasion. *Psychology, 57*, 345-374.

Creswell, J., & Plano Clark, V. (2010). *Designing and conducting mixed methods research*. Singapore: Sage.

Creswell, J., Plano Clark, V., Gutmann, M., & Hanson, W. (2003). Advanced mixed methods research designs. In A. Taskhakkori (Ed.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). London: Sage.

Cronbach, L. J. (1970). *Essentials of psychological testing*. New York: Harper & Row.

Curtin University. (2009). *Research ethics*. Retrieved 10 October, 2009, from www.research/curtin.edu.au

Darling-Hammond, L. (1998). Teachers and teaching: Testing hypotheses from a National Commission Report. *Educational Researcher, 27*, 5-15.

Darling-Hammond, L., & Sykes, G. (2003). Wanted: A national teacher supply policy for education: The right way to meet the "highly qualified teacher" challenge. *Educational Policy Analysis Archives, 11*, 1-55.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *Introduction: The discipline and practice of qualitative research*. Thousand Oaks, CA: Sage.

Depdiknas. (2003). Act of the Republic of Indonesia Number 20, Year 2003 on National Education System. Jakarta: Depdiknas Offset.

Depdiknas. (2003). *Undang-Undang Sistim Pendidikan Nasional Tahun 2003.* Retrieved 30 October, 2011, from www.depdiknas.go.id.

Dorman, J. P. (2003). Cross-national validation of the What Is Happening In this class (WIHIC) questionnaire using confirmatory analysis. *Learning Environments Research, 6*, 231-245.

Dorman, J. P., Adam, J. E., & Ferguson, J. M. (2002). Psychosocial environment and student self-handicapping in secondary school mathematics classes: A cross-national study. *Educational Psychology, 22*, 499-511.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology*. New York: Macmillan Library Reference USA.

Elmore, R. F. (2004). School reform from the inside out: Policy, practice and performance. Cambridge, MA: Harvard University Press.

Fan, X. (2001). Statistical significance and effect size in education research: Two sides of a coin. *The Journal of Educational Research, 94*, 275-282.

Farenga, S. J., & Joyce, B. A. (1998). Science-related attitudes and science course selection: A study of high-ability boys and girls. *Roeper Review, 20*, 247-251.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. *Advances in Experimental Social Psychology, 14*, 161-202.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers' College Record, 103*, 1013-1055.

Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE.

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

Fisher, D. L., & Fraser, B. J. (1983). Validity and use of classroom environment scale. *Educational Evaluation and Policy Analysis 5*, 261-271.

Fisher, D., Henderson, D., & Fraser, B. J. (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education, 25*, 125-133.

Fishman, B. J., Marx, R. W., Best, S., & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education, 19*, 643-658.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299.

Fraser, B. J. (1981). Using environmental assessments to make better classrooms. *Journal of Curriculum Studies, 13*, 131-144.

Fraser, B. J. (1989). Twenty years of classroom environment research: Progress and prospect. *Journal of Curriculum Studies, 21*, 307-327.

Fraser, B. J. (1990). Individualised Classroom Environment Questionnaire: Handbook and test master set. Melbourne: Australian Council for Educational Research.

Fraser, B. J. (1998). Development, validity and applications. *Learning Environments Research, 1*, 7-34.

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

Fraser, B. J. (2007). Classroom learning environments. In S.K. Abell & N.G. Lederman (Eds.), *Handbook of research on science education* (pp. 103-124). Mahwah, NJ: Lawrence Erlbaum.

Fraser, B. J. (2012). Classroom learning environments: Retrospect, context and prospect. In B. J. Fraser, K. G. Tobin and C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191–1239). New York: Springer.

Fraser, B. J., Aldridge, J. M., & Adolphe, F. S. G. (2010). A cross-national study of secondary science classroom environments in Australia and Indonesia. *Research in Science Education, 40*, 551-571.

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

Fraser, B. J., Aldridge, J. M., & Soerjaningsih, W. (2010). Instructor-student interpersonal interaction and student outcomes at the university level in Indonesia. *The Open Educational Journal, 3*, 21-33.

Fraser, B. J., & Butts, W. L. (1982). Relationship between perceived level of classroom individualization and science-related attitudes. *Journal of Research in Science Teaching, 19*, 143-154.

Fraser, B. J., & Fisher, D. L. (1983). Student achievement as a function of analysis person-environment fit: A regression surface. *British Journal of Educational Psychology, 53*, 88-99.

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Synthase of educational productivity research. *International Journal of Educational Research, 11*, 145-252.

Fraser, B. J., Fisher, D. L., & McRobbie, C. J. (1996, April). *Development, validation and use of personal and class forms of a new classroom environment instrument.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching, 32*, 399-422.

Fraser, B. J., & Lee, S. S. U. (2009). Science laboratory classroom environments in Korea. *Learning Environments Research, 12*, 67-84.

Fraser, B. J., & McRobbie, C. J. (1995). Science laboratory classroom environments at schools and universities: A cross-national study. *Educational Research and Evaluation, 1*, 289-317.

Fraser, B. J., McRobbie, C. J., & Giddings, G. J. (1993). Development and cross-national validation of a laboratory classroom environment instrument for senior high school science. *Science Education, 77*, 1-24.

Fraser, B. J., & O'Brien, P. (1985). Student and teacher perceptions of the environment of elementary school classrooms. *The Elementary School Journal, 85*, 567-580.

Fraser, B. J., & Treagust, D. F. (1986). Validity and use of an instrument for assessing classroom psychosocial environment in higher education. *Higher Education, 15,* 37-57.

Fraser, B. J., Treagust, D. F., & Dennis, N. C. (1986). Development of an instrument for assessing classroom psychosocial environment at universities and colleges. *Studies in Higher Education, 11*, 43-54.

Friedman, A., & Philips, M. (2004). Continuing professional development: Developing a vision. *Journal of Education and Work, 17*, 361-376.

Fullan, M. G. (2001). *The new meaning of educational change*. New York: Routledge.

Fullan, M. G. (2003). *Change forces with a vengeance*. London: Routledge Falmer.

Fullan, M. G., & Hargreaves, A. (2002). *Teacher development and educational change*. Philadelphia, PA: Routledge Falmer.

Fullan, M. G., Hill, P., & Crevola, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin Press.

Fullan, M. G., & Miles, M. B. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan, 73*, 745-752.

Fullan, M. G., & Stiegelbauer, S. (2007). *The new meaning of educational change*. New York: Teachers College Press.

Fuller, B. (1987). What factors raise achievement in the third world? *Review of Educational Research, 57*, 255-292.

Gaible, E., & Burns, M. (2005). Using technology to train teachers: Appropriate uses of ICT for teacher professional development in developing countries. Retrieved 28 March, 2012, from http://www.infodev.org/en/Publications.13.html

Gall, M. D., & Renchler, R. S. (1995). *Effective staff development for teachers: A research-based model*. Eugene, OR: Clearinghouse on Educational Management, University of Oregon.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective?: Results from a national sample of teachers. *American Educational Research Journal, 38*, 915-945.

Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research, 1*, 199-229.

Goh, S. C., Young, D. J., & Fraser, B. J. (1995). Pyschosocial climate and student outcomes in elementary mathematics classrooms. *Journal of Experimental Education, 64*, 29-40.

Goh, S. C., & Khine, M. S. (Eds.). (2002). Studies in educational learning environments: An international perspective. Singapore: World Scientific.

Goodall, J., Day, C., Lindsay, G., Muijs, D., & Harris, A. (2005). Evaluating the impact of Continuing Professional Development (CPD). London: DfES.

Griffin, G. A. (1983). Introduction: The work of staff development. In G. A. Griffin (Ed.), *Staff development. Eighty-second yearbook of The National Society for the Study of Education*. Chicago: University of Chicago Press.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough?: An experiment with data saturation and variability. *Field Methods, 18*, 59-82.

Guskey, T. R. (1986). Staff development and the process of teacher change. *American Educational Research Association, 15*(5), 5-12.

Guskey, T. R., (2000). *Evaluating professional development*. London, UK: Corwin Pr.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching, 8*, 381-391.

Guskey, T. R., & Huberman, M. (1995). *Professional development in education: New paradigms and practices*. New York: Teachers College Press.

Guskey, T. R., & Sparks, D. (2002). Linking professional development to improvements in student learning. Iowa: Kendall Hunt.

Guskey, T. R., & Yoon, K. S. (2009). What works in professional development? *Phi Delta Kappan, 90*, 495-500.

Hahs-Vaughn, D., Zygouris-Coe, V., & Fiedler, R. (2007). A hybrid evaluation model for evaluating online professional development. *Technology, Pedagogy and Education, 16*, 5-20.

Hanley, P., Maringe, F., & Ratcliffe, M. (2008). Evaluation of professional development: Deploying a process-focused model. *International Journal of Science Education, 30*, 711-725.

Harris, J. (2001). Logic model basics: After school at the YWCA of Asheville. *The Evaluation Exchange, 7*, 13-14.

Hassel, E. (1999). *Professional development: Learning from the best*. Oak Brook, IL: North Central Regional Educational Laboratory.

Hatton, E. J. (2000). Teachers' work as bricolage: Implications for teacher education. In S. J. Ball (Ed.), *British Journal of Sociology of Education* (pp. 337-357). London: Routledge-Falmer.

Henderson, D., Fisher, D. L., & Fraser, B. J. (2000). Interpersonal behaviour, learning environments and student outcomes in senior biology classes. *Journal of Research in Science Teaching, 37*, 26-43.

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences, 16*, 107-117.

Hirsh, S. (2006). Assessment inventory measures professional development quality. *National Staff Development Council, 27*, 63-63.

Hoban, G. F. (1997). Theories and models of professional development. In R. J. King, D. M. Hill & J. A. Retallick (Eds.), *Exploring professional development in education.* Riverwood, NSW: Social Science Press.

Hoban, G. F. (2002). *Teacher learning for educational change*. Buckingham: Open University Press.

Holloway, J. H. (2000). The promise and pitfalls of site-based management. *Educational Leadership, 57*, 81-87.

Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher, 17*, 10-16.

Huberman, M. (1995). Teacher development and instructional mastery. In A. Hargreaves & M. Fullan (Eds.), *Understanding teacher development* (pp. 122-142). London: Cassel Villiers House.

Hunt, D. E. (1975). Person-environment interaction: A challenge found waiting before it was tired. *Review of Educational Research, 45*, 219-230.

Indrajati, S. (2008, January). *Sinergi kurikulum dan ujian nasional.* Paper presented at the Sosialisasi Hasil Ujian Nasional, Semarang, Indonesia.

Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes and efficacy. *Education Policy Analysis Archives, 13*, 1-26.

Jacobs, J. K., Kawanaka, T., & Stigler, J. W. (1999). Integrating quantitative and qualitative approaches to the analysis of video data in classroom teaching. *International Journal of Educational Research, 31*, 717-724.

Jegede, O. J., Fraser, B. J., & Okebukola, P. A. (1994). Altering socio cultural beliefs hindering the learning of science. *Instructional Science, 22*, 137-152.

Jick, T. D. (1983). Mixing qualitative and quantitative research methods: Triangulation in action. In J. V. Maanen (Ed.), *Qualitative methodology* (pp. 135-148). Beverly Hills, CA: Sage.

Johnson, B., & McClure, R. (2004). Validity and reliability of a shortened, revised version of the Constructivist Learning Environment Survey (CLES). *Learning Environments Research, 7*, 65-80.

Johnson, D. W., & Johnson, R. T. (1995). Goal structures. In L. W. Anderson (Ed.), *International encyclopaedia of teaching and teacher education* (pp. 349-352). Tarrytown, NY: Pergamon

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*, 14-26.

Johnson, S. M. (2001). Can professional certification for teachers reshape teaching as a career?. *Phi Delta Kappan, 82*, 393-399.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: Association for Supervision & Curriculum Development.

Katharina, E. (2009). *English language teaching in Indonesia: Constraints on teachers' quality improvement.* Paper presented at the Applied Business Research & College Teaching and Learning Conference, Oahu, Hawaii, USA.

Kelly, P. (2006). What is teacher learning? A socio-cultural perspective. *Oxford Review of Education, 32*, 505-519.

Kennedy, C. (1998). Evaluation of management of change in ELT projects. *Applied Linguistics, 9*, 329-342.

Khine, M. S. (2002). Study of learning environment for improving science education in Brunei. In S. J. Goh & J. B. Kahle (Eds.), *In Studies in educational learning environments* (pp. 21-39). Singapore: World Scientific.

Khoo, H. S., & Fraser, B. J. (2008, March). Using classroom psychosocial environment in the evaluation of adult computer application courses in Singapore. *Technology, Pedagogy and Education, 17*, 67-81.

Killion, J. (2002). *Assessing impact: Evaluating staff development*. Oxford, OH: National Staff Development Council.

Kilpatrick, J. (2000). Reflections on verifying change in school mathematics. *Journal of Classroom Interaction*, *33*: 28-30.

Kim, H. B., Fisher, D. L., & Fraser, B. J. (2010). Classroom environment and teacher interpersonal behaviour in secondary science classes in Korea. *Evaluation & Research in Education 14*, 3-22.

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education, 29*, 871-893.

Kirkpatrick, D. (1976). Evaluation of training. In R. L. Craig & I. L. Bittell (Eds.), *Training and development handbook* (pp. 87-112). New York: McGraw-Hill.

Kirkpatrick, D., & Kirkpatrick, J. (2005). *Transferring learning to behavior: Using the four levels to improve performance*. San Francisco, CA: Berret-Koehler.

Klopfer, L. E. (1976). A structure for the affective domain in relation to science education. *Science Education, 60*, 299-312.

Koul, R. B., & Fisher, D. (2005). Cultural background and students' perceptions of science classroom learning environment and teacher interpersonal behaviour in Jammu, India. *Learning Environments Research, 8*, 195-211.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kuyper, H., van der Werf, M., & Lubbers, M. (2000). Motivation, meta-cognition and self-regulation as predictors of long-term educational attainment. *Educational Research and Evaluation, 6*, 181-205.

Kvale, S. (1996). Interviews: An introduction to qualitative research interviewing. Thousand Oaks, CA: Sage.

Kvale, S., & Brinkmann, S. (2009). Inter Views: Learning the craft of qualitative interviewing. Los Angeles: Sage.

Leary, T. (1957). *An interpersonal diagnosis of personality*. New York: Ronald Press Company.

Lee, S. S. U., Fraser, B. J., & Fisher, D. L. (2003). Teacher-student interactions in Korean high school science classrooms. *International Journal of Science and Mathematics Education, 1*, 67-85.

Lewin, K. (1936). Principles of topological psychology. New York: McGraw.

Lie, A. (2007). Education Policy and EFL Curriculum in Indonesia: Between the commitment to competence and the quest for higher test scores. *TEFLIN Journal, 18*, 1-14.

Lieberman, A. (1995). Practices that support professional development: Transforming conceptions of professional learning. *Phi Delta Kappan,* 76, 591-596.

Lieberman, A., & Miller, L. (2000). Teaching and teacher development: A new synthesis for a new century. In R. Brandt (Ed.), *Education in a New Era* (pp. 47-66). Alexandra, VA: Association for Supervision and Curriculum Development.

Lieberman, J., & Wilkins, E. (2006). The professional development pathways model: From policy to practice. *Kappa Delta Pi Record*, 42, 124-128.

Lightburn, M., & Fraser, B. J. (2007). Classroom environment and student outcomes among students using anthropometry activities in high school science. *Research in Science & Technological Education, 25*, 153-166.

Logan, K. A., Crump, B. J., & Rennie, L. J. (2006). Measuring the computer classroom environment: Lessons learned from using a new instrument *Learning Environments Research, 9*, 67-93.

Loucks-Horsley, S., Katherine, E., & Peter, S. W. (2003). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics, 99*, 258-271.

Love, N. (2002). Using data—getting results: Collaborative inquiry for school-based mathematics and science reform. Norwood, MA: Christopher-Gordon Publishers, Inc.

Lovitt, C., & Clarke, D. M. (1988). *Mathematics curriculum and teaching program*. Carlton, Victoria: Curriculum Corporation.

MacDowell-Goggin, M., & Fraser, B. J. (2004, April). *Effects of using a graphic organizer on science students' attitudes and classroom learning environments.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

MacLeod, C., & Fraser, B. J. (2010). Development, validation and application of a modified Arabic translation of the What Is Happening In this Class? (WIHIC) questionnaire. *Learning Environments Research, 13*, 1-21.

Mahady, R., Wardani, I. G. A. K., Irianto, B., Somerset, H. C. A., & Nielsen, D. (1996). *Secondary education in Indonesia: Strengthening teacher competency and student learning*. Jakarta, Indonesia: Ministry of Education and Culture.

Majeed, A. M., Fraser, B. J., & Aldridge, J. M. (2002). Learning environment and its association with student satisfaction among mathematics students in Brunei Darussalam. *Learning Environments Research, 5*, 203-226.

Maor, D., & Fraser, B. J. (1996). Use of classroom environment perceptions in evaluating inquiry-based computer assisted learning. *International Journal of Educational Research, 18*, 401-421.

Maor, D., & Fraser, B. J. (2005). An online questionnaire for evaluating students' and teachers' perceptions of constructivist multimedia in learning environments. *Research in Science Education, 35*, 221-244.

Margianti, E. S., Aldridge, J. M., & Fraser, B. J. (2002, April). *Investigating the learning environment and students' outcomes at the university level in Indonesia.* Paper presented at the American Educational Research Association, New Orleans, LA.

Martin-Dunlop, C., & Fraser, B. J. (2008). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. *International Journal of Science and Mathematics Education, 6*, 163-190.

Mathison, S. (1992). An evaluation model for in-service teacher education. *Evaluation and Program Planning, 15*, 255-261.

Mayadas, F., Bourne, J., & Moore, J. C. (2002). Introduction. In J. Bourne & J. C. Moore (Eds.), *Elements of quality online education* (pp. 7-12). MA: The Sloan Consortium.

McLaughlin, M. W., & Zarrow, J. (2001). Teachers engaged in evidence-based reform: Trajectories of teachers' inquiry, analysis and action. In A. Lieberman & L. Miller (Eds.), *Teachers caught in the action: Professional development that matters* (pp. 79-101). New York: Teachers College.

McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction.* New York, NY: Longman.

Mendiknas. (2006). Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006. Jakarta: Depdiknas.

Merriam, S. B. (2002). *Introduction to qualitative research*. San Francisco, CA: Jossey-Bass.

Middlewood, D., Parker, R., & Beere, J. (2005). *Creating a learning school*. London: Paul Chapman Publishing.

Mink, D. V., & Fraser, B. J. (2005). Evaluation of a K–5 mathematics program which integrates children's literature: Classroom environment and attitudes. *International Journal of Science and Mathematics Education, 3*, 59-85.

Mizel, H. (2003). Facilitator: 10; refreshments: 8; evaluation: 0. *Journal of the National Staff Development Council, 24*, 10-13.

Monahan, T. C. (1996). Do contemporary incentives and rewards perpetuate out-dated forms of professional development? *Journal of Staff Development, 17*, 44-47.

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

Moos, R. H. (1976). *The human context: Environmental determinants*. New York: John Wiley & Sons.

Moos, R. H. (1979). Evaluating educational environments: Procedures, measures, findings and policy implications. San Francisco, CA: Jossey Bass.

Moos, R. H. (1991). Connections between school, work and family settings. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environment: Evaluation, antecedents and consequences* (pp. 29-53). London: Pergamon.

Moos, R. H., & Trickett, E. J. (1987). *Classroom environment scale manual*. Palo Alto, CA: Consulting Psychologists Press.

Mueller, D. J. (1986). Measuring social attitudes: A handbook for researchers and practitioners. New York: Teacher College Press.

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

Na'im, M. (1995). The effect of the inquiry learning approach in biology on student achievement in a senior secondary school in Indonesia. Unpublished Doctoral Thesis, Curtin University of Technology, Perth.

Nair, C. S., & Fisher, D. L. (2001). Learning environments and student attitudes to science at the senior secondary and tertiary levels. *Issues in Educational Research, 11*, 12-31

Newmann, F. M., Rutter, R. A., & Smith, M. S. (1989). Organizational factors that affect school sense of efficacy, community and expectations. *Sociology of Education, 62*, 221-238.

Newman, I., Ridenour, C. S., Newman, C., & DeMarco, G. M. P. J. (2003). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 167-188). Thousand Oaks, CA: Sage.

Nix, R. K., Fraser, B. J., & Ledbetter, C. E. (2005). Evaluating an integrated science learning environment using the constructivist learning environment survey. *Learning Environments Research, 8*, 109-133.

Office for Standards in Education (Ofsted). (2004). Making a difference: The impact of award bearing in-service training on school improvement. London: Ofsted.

Ogbuehi, P. I., & Fraser, B. J. (2007). Learning environment, attitudes and conceptual development associated with innovative strategies in middle-school mathematics. *Learning Environments Research, 10*, 101-114.

Onwuegbuzie, A. J., & Leech, N. L. (2005). Taking the "Q" out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity: International Journal of Methodology, 39*, 267-298.

Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement. London: Pinter Publishers.

Osborne, J. (2003). Attitudes toward science: A review of the literature and its implications. *International Journal Science Education, 25*, 1049-1079.

Peiro, M. M., & Fraser, B. J. (2009). Assessment and investigation of science learning environments in the early childhood grades. In M. Ortiz & C. Rubio (Eds.), *Educational evaluation: 21st century issues and challenges* (pp.349-366). New York: Nova Publisher.

Pickett, L., & Fraser, B. J. (2009). Evaluation of a mentoring program for beginning teachers in terms of the learning environment and student outcomes in participants' school classrooms. In A. Selkirk & M. Tichenor (Eds.), *Teacher education: Policy, practice and research* (pp. 1-52). New York: Nova Publisher.

Purwadi, A., & Mulyoatmodjo, S. (2000). Education in Indonesia: Coping with challenges in the third millennium. *Journal of Southeast Asian Education, 3*, 79-102.

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

Raaflaub, C. A., & Fraser, B. J. (2002, April). *Investigating the learning environment in Canadian mathematics and science classrooms in which laptop computers are used.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Reid, N. (2006). Thoughts on attitude measurement. *Research in Science & Technological Education, 24*, 3-27.

Rentoul, A. J., & Fraser, B. J. (1979). Conceptualisation of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies, 11*, 233-245.

Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers: Strategies for teacher learning.* Cambridge: Cambridge University Press.

Richards, J. C., & Lockhart, C. (2000). *Reflective teaching in second language classrooms*. Cambridge: Cambridge University Press.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula, T. Buttery & E. Guyton (Eds.), *Handbook of research on teacher education* (pp. 102-119). New York: Simon & Schuster Macmillan.

Ritchie, J., & Lewis, J. (2003). Qualitative research practice: A guide for social science students and researchers. London: Sage.

Rochsantiningsih, D. (2006). Conducting action research in Indonesia. *Indonesian Journal of English Language Teaching, 2*, 21-35.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage.

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp.221-259). Thousand Oaks, CA: Sage.

Schopler, J., & Insko, C. A. (1992). The discontinuity effect in interpersonal and intergroup relations: Generality and mediation. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (pp. 121-151). New York: John Wiley.

Scott, R. H., & Fisher, D. L. (2004). Development, validation and application of a Malay translation of an elementary version of the Questionnaire on Teacher Interaction. *Research in Science Education, 34*, 173-194.

Scott Houston, L., Fraser, B. J., & Ledbetter, C. E. (2008). An evaluation of elementary school science kits in terms of classroom environment and student attitudes. *Journal of Elementary Science Education, 20*, 29-47.

Shah, J., Higgins, T., Friedman, & Ronald, S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology, 74*, 285-293.

Shawer, S. (2010). Classroom-level curriculum development: EFL teachers as curriculum developers, curriculum-makers and curriculum-transmitters. *Teaching and Teacher Education, 26*, 173-184.

Shulman, L. S., & Tamir, P. (1972). Research on teaching in the natural sciences. In R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp.1098-1148). Chicago, IL: Rand McNally.

Sink, C. A., & Spencer, L. R. (2005). My Class Inventory—Short Form as accountability tool for elementary school counsellors to measure classroom climate. *Professional School Counselling, 9*, 37-48.

Soerjaningsih, W., Aldridge, J. M., & Fraser, B. J. (2010). Achievement, satisfaction and learning environment among Indonesian computing students at the university level. *Open Education Journal, 3*, 21-33.

Soto-Rodriguez, M. T., & Fraser, B. J. (2004, April). A comparison of attitudes, achievement and classroom environment perceptions of LEP (Limited English Proficient) and non-LEP students in integrated science classrooms. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Sparks, D. (2002). Designing powerful professional development for teachers and principals. Oxford, OH: National Staff Development Council.

Sparks, D., & Hirsh, S. (2002). *A national plan for improving professional development*. Oxford, OH: National Staff Development Council.

Sparks, D., & Loucks-Horsley, S. (1990). Models of staff development. In W. R. Houston (Ed.), *Handbook of Research on Teacher Education* (pp. 234-250). New York: Macmillan.

Spinner, H., & Fraser, B. J. (2005). Evaluation of an innovative mathematics programme in terms of classroom environment, student attitudes, and conceptual development. *International Journal of Science and Mathematics, 3*, 267-293.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stern, G., Stein, M., & Bloom, B. (1956). *Methods in personality assessment*. Glencoe: The Free Press.

Stern, G. G. (1970). People in context: Measuring person-environment congruence in education and industry. New York: Wiley.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stiggins, R. J. (2001). The principal's leadership role in assessment. *NASSP Bulletin, 85*, 13-26.

Stigler, J., & Hiebert, J. (2004). *The teaching gap*. New York: The Free Press.

Supovitz, J. (2001). Translating teaching practice into improved student achievement. In S. Fuhrman (Ed.), From the capitol to the classroom: Standards-based reforms in the states. The one hundredth yearbook of the National Society for the Study of Education Part Two (pp. 81-98). Chicago: University of Chicago Press.

Supovitz, J., & Turner, H. (2000). The influence of standards-based reform on classroom practices and culture. *Journal of Research in Science Teaching, 37*, 1-18.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Taylor, P. C., Dawson, V., & Fraser, B. J. (1997, April). *Classroom learning environments under transformation: A constructivist perspective.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293-302.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: a typology with examples. *Journal of Mixed Methods Research, 1*, 77-99.

Teh, G., & Fraser, B. J. (1995a). Associations between student outcomes and geography classroom environment. *International Research in Geographical and Environmental Education, 4*, 3-18.

Teh, G., & Fraser, B. J. (1995b). Development and validation of an instrument for assessing the psychosocial environment of computer assisted learning. *Journal of Educational Computing Research, 12*, 177-193.

Teng, L. K., & Wong, A. F. L. (2009). Students' perceptions and attitudes in upper primary computer-assisted mathematics classroom. Retrieved 3 May, 2010, from www.aare.edu.au/01pap/won01433.htm

Thair, M., & Treagust, D. (2003). A brief history of a science teacher professional development initiative in Indonesia and the implications for centralised teacher development. *International Journal of Educational Development, 23*, 201-213.

Theobald, M. A. (2006). Increasing student motivation: Strategies for middle and high school teachers. Thousand Oaks, CA: Corwin.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher, 31*, 25-32.

Thompson, R. W. (2003). Assessing competencies of counsellors-in-training: Complexities in evaluating personal and professional development. *Counsellor Education and Supervision, 42*, 219-230.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33*, 529-554.

Tobin, K. G. (Ed.). (1993). *The practice of constructivism in science education*. Hillsdale, NJ: Lawrence Erlbaum & Associates.

Triandis, H. C. (1971). *Attitude and attitude change*. New York: Wiley.

Tribble, C. (2000). Designing evaluation into educational change processes. *ELT Journal, 54*(4), 319.

Umaedi. (2000). Manajemen peningkatan mutu berbasis sekolah. Jakarta: Depdikbud.

United Nations. (2010). *Principles and recommendations for population and housing census*. Retrieved 15 September, 2011, from http://unstats.un.org.

USAID. (2003). Teacher networks (MGMP) in junior secondary education in Indonesia. Jakarta: USAID Indonesia.

Velayutham, S., Aldridge, J. M. M., & Fraser, B. J. (2011). Development and validation of an instrument to measure students' motivation and self-regulation in science learning. *International Journal of Science Education*, 1, 1-21.

Wahyudi, & Treagust, D. (2004). The status of science classroom learning environments in Indonesian lower secondary schools. *Learning Environments Research, 7*, 43-63.

Walberg, H. J. (1981). A psychological theory of educational productivity. In F. Farley & N. Gordon (Eds.), *Psychology and education* (pp. 81-108). Berkeley, CA: McCutchan.

Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology, 59*, 414-419.

Waldrip, B. G., Fisher, D. L., & Dorman, J. P. (2009). Identifying exemplary science teachers through students' perceptions of their learning environment. *Learning Environments Research, 12*, 1-13.

Waldrip, B. G., Reene, P., Fisher, D. L., & Dorman, J. P. (2008). Changing primary students' perceptions of teacher interpersonal behaviours in science. *Research in Science Education, 38*, 213-235.

Walker, S. L. (2006). Development and validation of the Test of Geography-Related Attitudes (ToGRA). *Journal of Geography, 105*, 175-182.

Warren-Little, J. (2001). Teachers caught in the action: Professional development that matters. In A. Lieberman & L. Miller (Eds.), *On the school reform* (pp. 23-44). New York: Teachers College Press.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research, 79*, 702-739.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal Research of Science Teaching, 32*, 387-389.

Weiss, I. R., Pasley, J. D., Shimkus, E. S., & Smith, P. S. (2004). Looking inside the classroom: Science teaching in the United States. *Science Educator, 13*, 1-65.

Wiliam, D. (1998). A framework for thinking about research in mathematics and science education. In J. A. Malone, B. Atweh & J. R. Northfield (Eds.), *Research*

and supervision in mathematics and science education (pp. 1-16). Mahwah, NJ: Lawrence Erlbaum Associates.

Williams, R., & Burden, R. (2000). The role of evaluation in ELT project designs. *ELT Journal, 48*, 22-27.

Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education, 38*, 321-341.

Wong, A. F. L., & Fraser, B. J. (1997). A multilevel analysis of learning environments and student attitudes. *Educational Psychology, 17*, 449-468.

Wubbels, T. H., & Brekelmans, M. (1998). A comparison of student perceptions of Dutch physics teachers' interpersonal behaviour and their educational opinions in 1984 and 1993. *Journal of Research in Science Teaching, 34*, 447-466.

Wubbels, T. H., & Brekelmans, M. (2005). Two decades of research on teacher–student relationships in class. *International Journal of Educational Research, 43*, 6-24.

Wubbels, T.H., Brekelmans, M., & Hooymayers, H. P. (1992). Interpersonal teacher behaviour in the classroom. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 141-160). Oxford: Pergamon.

Wubbels, T., Creton, H. A., & Hooymayers, H. P. (1992). Review of research on teacher communication styles with use of the Leary model. *Journal of Classroom Interaction, 27*(1-12)

Wubbels, T. H., & Levy, J. (Eds.). (1993). Do you know what you look like: Interpersonal relationships in education. London: Falmer Press.

Yin, R. K. (2011). *Applications of case study research*. California, CA: Sage .

Yulaelawati, E. (2004). Kurikulum dan Pembelajaran: Filosofi, teori dan aplikasi. Bandung: Pakar Raya.

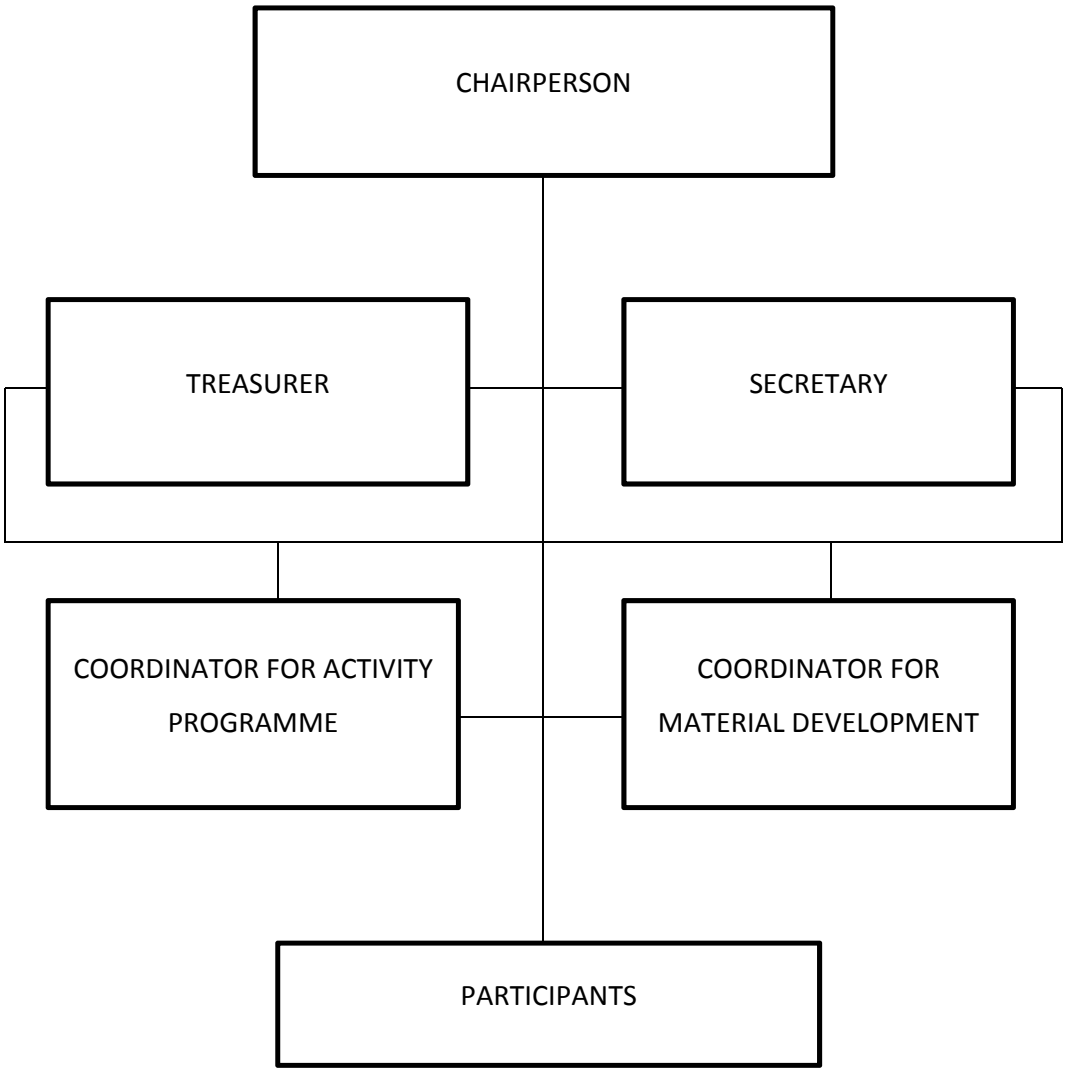
Zandvliet, D. B., & Fraser, B. J. (2005). Physical and psychosocial environments associated with networked classrooms. *Learning Environments Research, 8*, 1-17.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

Appendix A

Organisational Structure of the

MGMP Empowerment Programme


CHAIRPERSON
TREASURER
SECRETARY
COORDINATOR FOR ACTIVITY PROGRAMME
COORDINATOR FOR MATERIAL DEVELOPMENT
PARTICIPANTS

Appendix B

Example of Part of the Programme used during the MGMP Empowerment Programme

NO	CODE	PROGRAMME	TIME (in hours)
A	MU	GENERAL	
1	MU	Introduction to the Programme	2
	MU	Government Policy in Education	2
B	ING	BASIC	
1	ING 01	Theoretical Frameworks of Learning of English	4
2	ING 02	Curriculum Development and Lesson Planning	6
3	ING 03	Material Development	6
4	ING 04	School Based Curriculum	7
5	ING 05	IT Based Teaching Media	6
6	ING 06	Learning Strategies	6
7	ING 07	Teaching English Text	
	ING 08	Procedure	5
	ING 09	Descriptive	5
	ING 10	Narrative	5
	ING 11	Recount	5
	ING 12	Report	5
C	PEN	SUPPORTING MATERIALS	
	PEN 1	Lesson Study	8
	PEN 2	Classroom Action Research	8
		TOTAL	80

Appendix C

English Version of the Modified

What Is Happening In this Class?[1] (WIHIC) Questionnaire

[1] All scales in my study and included in this thesis were used with the permission of their authors.

What Is Happening In this Class?

Directions for Students

These questionnaires contain statements about practices which could take place in this class. You will be asked how often each practice takes place.

There are no 'right' or 'wrong' answers. Your opinion is what is wanted. Think about how well each statement describes what this class is like for you.

Draw a circle around

1	if the practice takes place	**Almost Never**
2	if the practice takes place	**Seldom**
3	if the practice takes place	**Sometimes**
4	if the practice takes place	**Often**
5	if the practice takes place	**Almost Always**

Be sure to give an answer for all questions. If you change your mind about an answer, just cross it out and circle another.

Some statements in this questionnaire are fairly similar to other statements. Don't worry about this. Simply give your opinion about all statements.

Practice Example
Suppose you were given the statement "I choose my partners for group discussion." You would need to decide whether you choose your partners 'Almost always', 'Often', 'Sometimes', 'Seldom' or 'Almost never'. If you selected 'Often' then you would circle the number 2 on your questionnaire.

STUDENT COHESIVENESS	Almost Never	Seldom	Some-times	Often	Almost Always
1. I make friendships among students in this class.	1	2	3	4	5
2. I know other students in this class.	1	2	3	4	5
3. I am friendly to members of this class.	1	2	3	4	5
4. Members of the class are my friends.	1	2	3	4	5
5. I work well with other class members.	1	2	3	4	5
6. I help other class members who are having trouble with their work.	1	2	3	4	5
7. Students in this class like me.	1	2	3	4	5
8. In this class, I get help from other students.	1	2	3	4	5
TEACHER SUPPORT	Almost Never	Seldom	Some-times	Often	Almost Always
9. The teacher takes a personal interest in me.	1	2	3	4	5
10. The teacher goes out of his/her way to help me.	1	2	3	4	5
11. The teacher considers my feelings.	1	2	3	4	5
12. The teacher helps me when I have trouble with the work.	1	2	3	4	5
13. The teacher talks with me.	1	2	3	4	5
14. The teacher is interested in my problems.	1	2	3	4	5
15. The teacher moves about the class to talk with me.	1	2	3	4	5
16. The teacher's questions help me to understand.	1	2	3	4	5
INVOLVEMENT	Almost Never	Seldom	Some-times	Often	Almost Always
17. I discuss ideas in class.	1	2	3	4	5
18. I give my opinions during class discussions.	1	2	3	4	5
19. The teacher asks me questions.	1	2	3	4	5
20. My ideas and suggestions are used during classroom discussions.	1	2	3	4	5
21. I ask the teacher questions.	1	2	3	4	5
22. I explain my ideas to other students.	1	2	3	4	5
23. Students discuss with me how to go about solving problems.	1	2	3	4	5
24. I am asked to explain how I solve problems.	1	2	3	4	5

FINDING REFERENCE	Almost Never	Seldom	Some-times	Often	Almost Always
25. I look for references to test my ideas.	1	2	3	4	5
26. I am asked to find references for statements.	1	2	3	4	5
27. I look for references to answer questions coming from discussions.	1	2	3	4	5
28. I explain the meaning of statements	1	2	3	4	5
29. I look for references to answer questions which puzzle me.	1	2	3	4	5
30. I look for references to answer the teacher's questions.	1	2	3	4	5
31. I find out answers to questions by looking for references	1	2	3	4	5
32. I solve problems by using information obtained from references.	1	2	3	4	5
TASK ORIENTATION	Almost Never	Seldom	Some-times	Often	Almost Always
33. Getting a certain amount of work done is important to me.	1	2	3	4	5
34. I do as much as I set out to do.	1	2	3	4	5
35. I know the goals for this class.	1	2	3	4	5
36. I am ready to start this class on time.	1	2	3	4	5
37. I know what I am trying to accomplish in this class.	1	2	3	4	5
38. I pay attention during this class.	1	2	3	4	5
39. I try to understand the work in this class.	1	2	3	4	5
40. I know how much work I have to do.	1	2	3	4	5

COOPERATION	Almost Never	Seldom	Some-times	Often	Almost Always
41. I cooperate with other students when doing assignment work.	1	2	3	4	5
42. I share my books and resources with other students when doing assignments.	1	2	3	4	5
43. When I work in groups in this class, there is teamwork.	1	2	3	4	5
44. I work with other students on projects in this class.	1	2	3	4	5
45. I learn from other students in this class.	1	2	3	4	5
46. I work with other students in this class.	1	2	3	4	5
47. I cooperate with other students on class activities.	1	2	3	4	5
48. Students work with me to achieve class goals.	1	2	3	4	5
EQUITY	Almost Never	Seldom	Some-times	Often	Almost Always
49. The teacher gives as much attention to my questions as to other students' questions.	1	2	3	4	5
50. I get the same amount of help from the teacher as do other students.	1	2	3	4	5
51. I have the same amount of say in this class as other students.	1	2	3	4	5
52. I am treated the same as other students in this class.	1	2	3	4	5
53. I receive the same encouragement from the teacher as other students do.	1	2	3	4	5
54. I get the same opportunity to contribute to class discussions as other students.	1	2	3	4	5
55. My work receives as much praise as other students' work.	1	2	3	4	5
56. I get the same opportunity to answer questions as other students.	1	2	3	4	5

Appendix D

English Version of the Enjoyment of

Enjoyment of English Classes[2] Attitude Scale

[2] The scale was used in my study with the permission of the authors.

Enjoyment of English Classes

Directions for Students

These questionnaires contain statements about practices which could take place in this class. You will be asked how often each practice takes place.

There are no 'right' or 'wrong' answers. Your opinion is what is wanted. Think about how well each statement describes what this class is like for you.

Draw a circle around

1	if the practice takes place	**Almost Never**
2	if the practice takes place	**Seldom**
3	if the practice takes place	**Sometimes**
4	if the practice takes place	**Often**
5	if the practice takes place	**Almost Always**

Be sure to give an answer for all questions. If you change your mind about an answer, just cross it out and circle another.

ENJOYMENT OF ENGLISH CLASSES		Almost Never	Seldom	Some-times	Often	Almost Always
1.	I look forward to English lessons.	1	2	3	4	5
2.	English lessons are fun.	1	2	3	4	5
3.	I enjoy the activities we do in English classes.	1	2	3	4	5
4.	English is one of the most interesting school subjects.	1	2	3	4	5
5.	I prefer school have more English lessons each week.	1	2	3	4	5
6.	I enjoy going to English lessons.	1	2	3	4	5
7.	I would like to belong to an English conversation club.	1	2	3	4	5
8.	I like talking to my friends in English.	1	2	3	4	5
9.	The material covered in English lessons is interesting.	1	2	3	4	5
10.	I feel satisfied after an English lesson.	1	2	3	4	5

Appendix E

Indonesian Version of the
Modified What Is Happening In this Class? (WIHIC)
Questionnaire

Apa Yang Sedang Terjadi di Kelas Ini?

Petunjuk pengisian angket

Angket berikut ini berisi pernyataan- pernyataan tentang kegiatan/ kejadian-kejadian yang (mungkin) terjadi di kelas pada saat pelajaran Bahasa Inggris berlangsung. Anda akan ditanya berapa sering kegiatan/ kejadian tersebut terjadi.

Dalam pengisian angket ini tidak ada jawaban benar atau salah. Hanya pendapat anda yang kami inginkan. Pikirkan seperti apa keadaan kelas anda pada saat pelajaran Bahasa Inggris berlangsung.

Lingkarilah:

1 jika pernyataan/ kejadian tersebut **hampir tidak pernah terjadi**

2 jika pernyataan/ kejadian tersebut **jarang terjadi**

3 jika pernyataan/ kejadian tersebut **kadang- kadang terjadi**

4 jika pernyataan/ kejadian tersebut **sering terjadi**

5 jika pernyataan/ kejadian tersebut **hampir selalu terjadi**

Jawab semua pertanyaan. Jika anda akan mengganti jawaban, silang jawaban tersebut dan lingkari jawaban yang lain.

Nama Siswa : …….

Jenis Kelamin : ……

Nama Guru : …….

Sekolah / Kelas : …….

KEDEKATAN SISWA		Hampir Tidak Pernah	Jarang	Kadang Kadang	Sering	Hampir Selalu
1.	Saya berteman dengan semua siswa di kelas ini.	1	2	3	4	5
2.	Saya kenal semua siswa di kelas ini.	1	2	3	4	5
3.	Saya ramah dengan semua siswa di kelas ini.	1	2	3	4	5
4.	Semua siswa di kelas ini adalah teman saya.	1	2	3	4	5
5.	Saya bekerjasama dengan baik dengan siswa lain di kelas ini.	1	2	3	4	5
6.	Saya membantu siswa lain di kelas ini yang mempunyai kesulitan dalam mengerjakan tugasnya.	1	2	3	4	5
7.	Siswa- siswa di kelas ini menyukai saya.	1	2	3	4	5
8.	Di kelas ini, saya mendapat bantuan dari siswa lainnya.	1	2	3	4	5

DUKUNGAN GURU		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
9.	Guru Bahasa Inggris memperhatikan saya.	1	2	3	4	5
10.	Guru Bahasa Inggris berusaha membantu saya.	1	2	3	4	5
11.	Guru Bahasa Inggris menghargai perasaan-perasaan saya.	1	2	3	4	5
12.	Guru Bahasa Inggris membantu saya ketika mengalami kesulitan dalam mengerjakan tugas	1	2	3	4	5
13.	Guru Bahasa Inggris mengajak saya berbicara.	1	2	3	4	5
14.	Guru Bahasa Inggris memperhatikan masalah saya.	1	2	3	4	5
15.	Guru Bahasa Inggris memotivasi teman-teman agar berbicara dengan saya.	1	2	3	4	5
16.	Pertanyaan- pertanyaan guru Bahasa Inggris membantu saya untuk memahami pelajarannya.	1	2	3	4	5

KETERLIBATAN SISWA		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
17.	Saya berdiskusi berbagai hal dalam pelajaran Bahasa Inggris.	1	2	3	4	5
18.	Saya berpendapat selama diskusi dalam pelajaran Bahasa inggris.	1	2	3	4	5
19.	Guru Bahasa Inggris bertanya kepada saya.	1	2	3	4	5
20.	Ide dan saran- saran saya diterima selama diskusi dalam pelajaran Bahasa Inggris	1	2	3	4	5
21.	Saya bertanya kepada guru selama pelajaran Bahasa Inggris.	1	2	3	4	5
22.	Saya menjelaskan ide- ide saya kepada teman- teman.	1	2	3	4	5
23.	Saya berdiskusi dengan teman- teman ketika menjawab soal- soal dalam pelajaran Bahasa Inggris	1	2	3	4	5
24.	Saya diminta untuk menjelaskan bagaimana cara saya menyelesaikan soal-soal dalam pelajaran Bahasa Inggris.	1	2	3	4	5

MENCARI REFERENSI		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
25.	Dalam pelajaran Bahasa Inggris, saya mencari referensi untuk menguji ide- ide saya.	1	2	3	4	5
26.	Dalam pelajaran Bahasa Inggris, saya diminta mencari referensi untuk mendukung pernyataan- pernyataan / jawaban.	1	2	3	4	5
27.	Dalam pelajaran Bahasa Inggris, saya mencari referensi untuk menjawab pertanyaan- pertanyaan yang muncul dalam diskusi.	1	2	3	4	5
28.	Dalam pelajaran Bahasa Inggris, saya menjelaskan maksud pernyataan-pernyataan/jawaban.	1	2	3	4	5
29.	Dalam pelajaran Bahasa Inggris, saya mencari referensi untuk menjawab pertanyaan yang membingungkan saya.	1	2	3	4	5
30.	Dalam pelajaran Bahasa Inggris, saya mencari referensi untuk menjawab pertanyaan guru.	1	2	3	4	5
31.	Dalam pelajaran Bahasa inggris, saya mendapatkan jawaban dari hasil pencarian referensi.	1	2	3	4	5
32.	Dalam pelajaran Bahasa Inggris, saya mengerjakan soal- soal dengan menggunakan informasi yang saya dapatkan dari referensi.	1	2	3	4	5
ORIENTASI TUGAS		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
33.	Berhasil menyelesaikan tugas pelajaran Bahasa Inggris merupakan hal yang penting bagi saya.	1	2	3	4	5
34.	Saya mengerjakan tugas pelajaran Bahasa Inggris seperti yang diperintahkan guru.	1	2	3	4	5
35.	Saya tahu tujuan belajar Bahasa Inggris.	1	2	3	4	5
36.	Saya siap mengikuti pelajaran Bahasa Inggris tepat waktu.	1	2	3	4	5
37.	Saya tahu apa yang harus saya capai dalam pelajaran Bahasa Inggris.	1	2	3	4	5
38.	Saya memperhatikan pelajaran Bahasa Inggris secara penuh.	1	2	3	4	5
39.	Saya berusaha memahami tugas- tugas dalam pelajaran Bahasa Inggris.	1	2	3	4	5
40.	Saya tahu berapa banyak tugas yang harus saya kerjakan/ selesaikan dalam pelajaran Bahasa Inggris.	1	2	3	4	5

KERJASAMA		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
41.	Saya bekerjasama dengan siswa- siswa lain dalam mengerjakan tugas Bahasa Inggris.	1	2	3	4	5
42.	Saya memakai buku dan sarana lain bersama teman- teman dalam mengerjakan tugas Bahasa Inggris.	1	2	3	4	5
43.	Ketika kerja kelompok, saya bekerjasama dengan baik dengan teman- teman.	1	2	3	4	5
44.	Saya bekerja dengan siswa- siswa lain ketika menyelesaikan tugas kelompok (project) Bahasa Inggris.	1	2	3	4	5
45.	Dalam pelajaran Bahasa Inggris, saya belajar dari siswa lain.	1	2	3	4	5
46.	Saya bekerja dengan siswa lain dalam pelajaran Bahasa Inggris.	1	2	3	4	5
47.	Saya bekerja dengan siswa lain dalam kegiatan belajar Bahasa Inggris.	1	2	3	4	5
48.	Siswa- siswa lain bekerjasama dengan saya untuk mencapai tujuan pelajaran Bahasa Inggris.	1	2	3	4	5
	PERSAMAAN	Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
49.	Guru Bahasa Inggris saya memberikan perhatian yang sama terhadap pertanyaan saya dan teman- teman.	1	2	3	4	5
50.	Guru Bahasa Inggris saya memberikan bantuan yang sama kepada saya dan teman- teman saya.	1	2	3	4	5
51.	Guru Bahasa Inggris saya memberikan kesempatan berbicara yang sama kepada saya dan teman- teman saya.	1	2	3	4	5
52.	Saya diperlakukan sama seperti teman-teman saya dalam pelajaran Bahasa Inggris.	1	2	3	4	5
53.	Guru Bahasa Inggris saya memberikan motivasi yang sama kepada saya dan teman- teman saya.	1	2	3	4	5
54.	Saya mendapatkan kesempatan yang sama seperti teman-teman saya ketika berdiskusi dalam Bahasa Inggris.	1	2	3	4	5
55.	Guru Bahasa Inggris saya memberikan penghargaan yang sama kepada saya dan teman- teman saya.	1	2	3	4	5
56.	Guru Bahasa Inggris memberikan kesempatan yang sama kepada saya dan teman- teman dalam menjawab pertanyaan.	1	2	3	4	5

Appendix F

Indonesian Version of the

Enjoyment of English Classes Attitude Scale

Skala Sikap

Petunjuk pengisian angket

Angket berikut ini berisi pernyataan- pernyataan tentang kegiatan/ kejadian-kejadian yang (mungkin) terjadi di kelas pada saat pelajaran Bahasa Inggris berlangsung. Anda akan ditanya berapa sering kegiatan/ kejadian tersebut terjadi. Dalam pengisian angket ini tidak ada jawaban benar atau salah. Hanya pendapat anda yang kami inginkan. Pikirkan seperti apa keadaan kelas anda pada saat pelajaran Bahasa Inggris berlangsung.

Lingkarilah:

1 jika pernyataan/ kejadian tersebut **hampir tidak pernah terjadi**

2 jika pernyataan/ kejadian tersebut **jarang terjadi**

3 jika pernyataan/ kejadian tersebut **kadang- kadang terjadi**

4 jika pernyataan/ kejadian tersebut **sering terjadi**

5 jika pernyataan/ kejadian tersebut **hampir selalu terjadi**

Jawab semua pertanyaan. Jika anda akan mengganti jawaban, silang **(X)** jawaban tersebut dan lingkari jawaban yang lain. Semua pendapat dan jawaban bersifat rahasia dan tidak akan dipublikasikan. Semua jawaban hanya digunakan untuk keperluan penelitian.

Nama Siswa : .....

Jenis Kelamin :.....

Nama Guru : .....

Sekolah / Kelas : .....

SIKAP		Hampir Tidak Pernah	Jarang	Kadang kadang	Sering	Hampir Selalu
57.	Saya mengharapkan saat pelajaran Bahasa Inggris tiba.	1	2	3	4	5
58.	Pelajaran Bahasa Inggris menyenangkan.	1	2	3	4	5
59.	Saya menyukai kegiatan yang dilakukan dalam pelajaran Bahasa Inggris.	1	2	3	4	5
60.	Bahasa Inggris merupakan salah satu pelajaran yang paling menyenangkan.	1	2	3	4	5
61.	Saya berpendapat/ berpikir agar sekolah sebaiknya menambah jam pelajaran Bahasa Inggris setiap minggunya.	1	2	3	4	5
62.	Saya senang mengikuti pelajaran Bahasa Inggris.	1	2	3	4	5
63.	Saya mengikuti kelompok percakapan Bahasa Inggris.	1	2	3	4	5
64.	Saya berbicara dengan teman dalam Bahasa Inggris.	1	2	3	4	5
65.	Materi pelajaran Bahasa Inggris menarik.	1	2	3	4	5
66.	Saya merasa puas dan senang setelah mengikuti pelajaran Bahasa Inggris.	1	2	3	4	5

Appendix G

Response Evaluation –
Questions to Guide the Teachers'
Reflective Journal Writing

Response Evaluation[3]

Teacher's Reflective Journal

Name : _____

Date : _____

Write your feeling about the workshops provided today in 10 to 15 minutes. The followings are questions that may guide you in writing your reflection.

1. Was the programme in general satisfactory?
2. Were the programme objectives clear, attainable and measurable?
3. Was time management properly observed?
4. Was the programme successful in conveying new knowledge about approaches to teaching English?
5. Did the programme deepen my understanding of how English is acquired?
6. Was the programme a useful forum for exchanging and developing ideas?
7. Was the programme relevant and applicable to my teaching situation?
8. Will the programme positively influence my teaching in the future?
9. Was the programme successful in introducing me to new ideas for the English class?
10. Were the handouts informative and useful? Was the quality of the presenter satisfactory?

[3] Source of questions

Response Evaluation Form (Guskey, 2000)

Appendix H

Classroom Observation Checklist

CLASSROOM OBSERVATION CHECKLIST

TEACHER: _____ SUBJECT: _____

ROOM: _____ PERIOD: _____

DATE: _____

ENGAGING AND SUPPORTING STUDENTS

N/A	NO	YES	ACTIVITIES
			1. Connects students' prior knowledge, life experience and interests with learning goals. 2. Uses a variety of instructional strategies and resources to respond to students' diverse needs. 3. Facilitates learning experiences that promote autonomy, interaction and choice. 4. Engages students in problem solving, critical thinking and other activities that make subject matter meaningful. 5. Promotes self-directed, reflective learning for all students

UNDERSTANDING AND ORGANISING SUBJECT MATTER

N/A	NO	YES	ACTIVITIES
			1. Demonstrates knowledge of subject matter content and student development. 2. Organises curriculum to support student' understanding of subject matter. 3. Interrelates ideas and information within across subject matter areas. 4. Develops student' understanding through instructional strategies that are appropriate to the subject matter. 5. Uses materials, resources and technologies to make subject matter accessible to students.

PLANNING INSTRUCTION AND DESIGNING LEARNING EXPERIENCES FOR ALL

N/A	NO	YES	ACTIVITIES
			1. Draws on and values students' backgrounds, interest and development learning needs. 2. Establishes and articulates goals for student learning. 3. Develops and sequences instructional activities and materials for student learning. 4. Designs long-term and short-term plans for student learning. 5. Modifies instructional plans to adjust for student needs.

This checklist is translated from the Classroom Observation Checklist used in the study.

ASPECTS OF LEARNING ENVIRONMENT SURVEY

Description of activities		3 minute intervals																		Total	
1	Student helps other class members who are having trouble with their work																				SC
2	Students work well with each other																				
3	Teacher moves around the class to talk with the students																				TS
4	Teacher praises or encourages student's action or behaviour																				
5	Teacher accepts or uses students' ideas																				
6	Teacher answer students' questions																				
7	Teacher asks questions																				IV
8	Students ask questions																				
9	Students respond to teachers questions or orders																				
10	Students explain their ideas to the class																				
11	Students use references (books, dictionaries or internet)																				FR
12	Teacher supports students' ideas of using references																				
13	Teacher explains the goal of the lesson																				TO
14	Students pay attention to the teacher																				
15	Students are engaged during the lesson																				
16	Students work in groups																				COP

17	Students cooperate with other students in this class																					
18	Students work well with the members of the group																					
19	Teacher gives the same opportunities to students to answer questions																					EQ
20	Teacher ignores students' answers																					
21	Students are interested in the English lesson																					ATT
22	Students look happy when engaged in English lesson																					
23	Students look satisfied after the English lesson																					

Appendix I

Example Interview Schedule - Teachers

Interview Schedule – Teachers

The interviews aimed to scrutinise the teachers' views of the teacher professional development programme and to ask their opinion about whether there were any contextual factors that might promote or impede the implementation. Sample interview items for each purpose are provided below.

Part A

The following questions will guide the interviews to probe the teachers' attitudes towards teacher professional development experience and their views on whether the programme has expanded their knowledge and skills.

1. What do you think about the MGMP Empowerment programme that you are participating in?
2. Do you think the suggestions delivered in the programme were compatible with your classroom? Why? or Why not?
3. In what ways do you think the information might be of value to you?
4. Do you think you will find the ideas useful when you go back to your classroom?

Part B

The following questions were used in the interviews with teachers to expose their opinions regarding the contextual factors that promote or impede the implementation.

1. In what ways did the MGMP Empowerment programme impact your classroom?
2. Following the programme, in what ways are your classrooms different?
3. How do you think your teaching changed as a result of the programme?
4. How did your students react to these changes in your teaching practices?
5. Do you think that your teaching changed as a result of your participation in the MGMP Empowerment professional development?
6. Do you get enough support from your principal and superintendent in the implementation of the professional development?