CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

This chapter discusses the relevant research related to the topic of this study and tries to address the gap in knowledge on issues pertaining to teacher evaluation. The discussion covers several aspects such as, instructional practice, definition of teacher evaluation, the need for teacher evaluation, formative and summative evaluation, rationale for evaluating teachers, teacher evaluation methods, instruments and criteria used for evaluation, as well as utilization of evaluation findings. This is followed by a discussion on the administrator's role in teacher evaluation, challenges faced in implementing teacher evaluation practices, models of teacher evaluation and an overview of teacher evaluation practices in Malaysia. This chapter concludes with a review of empirical studies on teacher evaluation practices in Malaysia and elsewhere.

Instructional Practice

Instructional practice refers to the teacher's core business in the classroom, that is, to impart knowledge or skills to all students (Smith, 1987). In other words, through the process of teaching, students are taught how to acquire knowledge, skills, values and other abilities. Teachers are in constant contact with students under their care and therefore they can influence how children interpret the knowledge given to them. They also have the monumental task of instilling in their students the skills or ability of applying the knowledge they have received when they leave the school environment.

Most pedagogical literature indicates that the concept of learning is usually intertwined with teaching, and thus the use of the term 'teaching-learning'. This signifies that teaching involves learning and thus, the notion that, when a teacher teaches in a

classroom, students should be able to learn something (Smith, 1987). While this might not be true in all cases, most effective teaching constitutes both teaching and learning.

In the literature we find that the question of effective teaching is at the core of most teacher evaluation systems, although no one can actually explain what is meant by the term effective or good teaching (Burgess, 1989; Darling-Hammond, 1990; Stronge 2007). Most of the time, the teacher evaluation system is implemented to produce evidence that effective teaching and learning is taking place in the classroom. Teacher quality is another term that is not explicitly defined but it is also implicated in the rationale for carrying out teacher evaluation. According to Elliot (1989, p.82) "the quality of teaching cannot be assessed in terms of performance-referenced criteria, but only in terms of the personal qualities displayed in the performance".

The criteria for evaluating teachers generally look at how teachers perform their teaching tasks to ensure that intended learning outcomes are achieved. This will encompass the full range of planning, instruction, assessment, and reflection stages of teaching. Teacher performance in all these stages determines whether an effective outcome is achieved. The items in the evaluation instruments stress excellence and high quality in teaching, but most evaluators have very little knowledge of how to identify these qualities or to decide what constitutes incompetence among teachers (Raths & Lyman, 2003). Most people say it is difficult to measure teachers' work (Burgess, 1989) because teaching could be seen as a labor, a craft, a professional activity and as an art (Elliot, 1989). Nevertheless, whatever the conception of teaching, teacher performance is being continuously assessed based on what the teacher does to ensure that learners acquire knowledge, skills, beliefs and attitude.

The implementation of a teacher evaluation system which includes classroom observation, interview, peer review, portfolio assessment, self-appraisal, student test scores, student survey and parent survey will help gather data on teacher performance.

This data can be used to improve instructional practice in the classroom. But Peterson (2000) states that the current teacher evaluation system makes very little contribution towards recognizing innovative teaching which can be adopted by teachers to improve instructional practice in the classroom.

Defining Teacher Evaluation

A review of the literature indicates that various terms are used to describe the act of estimating or judging the nature or value of teachers' performance. The diverse terms used include teacher evaluation, teacher appraisal, personnel review, staff assessment, personnel appraisal, efficiency rating, performance evaluation, staff ratings, and merit rating. In view of this, some clarification is required on terminology to ensure that there is no ambiguity in the use of the terms to refer to the process of judging or assessing teacher performance in this study. In particular, some distinctions must be made between teacher evaluation and teacher appraisal, as well as formative and summative evaluation.

In the United States, the term teacher evaluation refers to "assessments of the qualifications, competence, or performance of individual teachers for licensing, selection, continuation, promotion, tenure, professional growth, merit pay or national certification" (Dwyer & Stufflebeam, 1996, p. 769). Darling-Hammond (1990, p. 20) has defined evaluation as the process of "collecting and using information to judge the worth of something". She further explains that teacher evaluation includes the practice of collecting data and judging the worth, effectiveness and appropriateness of instructional practice in the classroom.

As for the British context, government documents cited in Burgess (1989, p. 26) explain that "staff appraisal involves qualitative judgments about performance and, although it may start as self-appraisal by the teacher, it will normally involve judgments by other persons responsible for that teacher's work." Holly (1989) defines appraisal as

'the forming of qualitative judgments" and therefore considers it crucial to teaching and professional development. According to Elliot (1989) the term 'appraisal' indicates judgment on teacher performance and 'evaluation' involves a broader scope in the educational setting. Stake (1989, p.13) looks at teacher evaluation as the "appraisal of qualification and performance of the individual teacher."

It is interesting to note that though various terms have been used in the literature to denote teacher evaluation, all these terms look at some common aspects of the teaching profession, such as competence, knowledge, qualifications, skills and professionalism of teachers. The process of evaluating teachers applies to all teachers who contribute towards the enhancement of instructional practice and not to a few selected ineffective teachers.

For the purpose of this study, the term 'teacher evaluation' is used to denote the systematic evaluation of a teacher's performance in relation to his/her defined professional roles and responsibilities in the Malaysian national primary schools. The investigation encompassed both formative and summative teacher evaluation practices in the selected schools. Formative evaluation practices were carried out throughout the year to improve instructional practice and promote teacher development, whereas summative evaluation was performed annually to give rewards and for personnel decision-making.

The Need for Teacher Evaluation

Appraising teaching is considered an integral part of the practice of evaluating the school itself (Stake, 1989). Teachers, who are professional educators involved in establishing a learning environment in the school must be evaluated like any other professionals in an organization. Though the teacher evaluation process is most of the time considered an intrusion in most schools, it has to be carried out for the improvement of educational practice and building confidence in teacher quality (Smith, 2001). In order

for this to happen school administrators must be constantly reminded that teacher evaluation should be focused on continuing professional development and not merely identifying teacher incompetence (Darling-Hammond, 1990).

Though there are numerous problems associated with defining and measuring teacher performance, there is still a need to carry out teacher evaluation to check the progress made by teachers and give feedback for enhancement and professional growth (Dwyer & Stufflebeam, 1996). Stake (1989) asserts that in order to evaluate teaching effectively and to ensure the validity of its findings, other factors need to be considered. These include "institutional goals, classroom environments, administrative organization and operations, curricular content, student achievement and the impact of school programmes on the community" (p.13). In the past, not much resources and organizational commitment were given to teacher evaluation and this had led to it having very little effect on matters pertaining to teaching (Darling-Hammond, 1990). Currently, the trend has changed towards using evaluation results to make important decisions about teachers and their work in schools.

When there is a better understanding of the overall teacher evaluation process, it can be "designed to support teaching and professional development, enhance personal and collaborative enquiry, promote critique, and contribute to an evolving pedagogy" (Holly, 1989 p.100). Intervention strategies based on evaluation will provide avenues to eventually improve teacher practice and pave the way towards more professionalism in teaching. But the question that is frequently asked is how school administrators utilize the information obtained from evaluation to improve teacher performance?

Research indicates that in order to be able use the results of an evaluation, the process must be credible, fitting for its purpose and feasible (Darling-Hammond, 1990). If these criteria are met, then school administrators can safely use their evaluation findings to improve instructional practice, plan professional development programs and

make personnel decisions. In order to obtain comprehensive data on teacher performance in school, administrators must take into account both the formative and summative evaluation findings.

Formative and Summative Evaluation

There are two types of evaluation, that is, formative evaluation which is designed to improve practice, and summative evaluation which provides information for making personnel decisions such as tenure, merit pay and promotion (Scriven, 1980). It is generally believed that through formative evaluation the need to improve weak areas can be emphasized. This form of evaluation enables administrators to make judgments about the work accomplished; to ascertain a teacher's knowledge, attitudes and skills, and to comprehend the changes that have taken place in these; and to enhance their ability to appraise their learning and performance (Smith, 2001). Educationists agree that for true evaluation to occur and improvements to be made, emphasis must be placed on formative evaluation.

When summative evaluation is conducted it allows organizations to show that the program or project has met its intended goals or to indicate that teachers or personnel have attained the requisite standard (Smith, 2001). This form of evaluation becomes a checkpoint when decisions need to be made about pay, promotion or release (Hunter, 1988b). Often, results of summative evaluation can determine if one is retained, dismissed or promoted (Peterson, 2004). This is not the case in the Malaysian context. Teachers are hardly retained or dismissed based on the evaluation findings. But in some instances it is used for promotional purposes.

Smith (2001) says that summative evaluation is considered a public process as it caters to the needs of the administrators, whereas formative evaluation is considered more confidential since it should be restricted to the teacher's consumption (Smith,

2001). Both these types of evaluation are usually implemented sequentially and not simultaneously (Hunter, 1988a), and there is a clear distinction of purpose for carrying out the respective type of evaluation (Peterson, 2004).

The Rationale for Evaluating Teachers

There is a need for a greater understanding of the phenomenon of evaluating teacher competency, assessing teacher quality and improving practice. The underlying rationale for evaluating teachers will determine the models that will be used and the instruments to gather the required data. It is generally believed that the purpose of teacher evaluation is for improvement in teaching performance (Duke, 1995), but there can be many other rationales for the implementation of teacher evaluation in schools.

Dwyer and Stufflebeam (1996) have stated several reasons for conducting teacher evaluation, and these include "preparing teachers, licensing teachers, selecting and monitoring teachers as well as fostering and certifying the professionalisation of teachers" (p.773). They have also explained how teacher evaluation has been used to remove incompetent teachers and to create a career ladder program that offers financial and other forms of inducement to encourage professional development or for accomplishing improved student learning.

Scriven (1990, p. 78) explains that teacher evaluation is carried out to make decisions about the 'merit', 'worth' and 'value' of teachers. 'Merit' indicates the extent to which a teacher meets the accepted standards of the teaching profession and 'worth' is the value of the teacher to the school as a learning organization (Scriven, 1990). These are the three common terms used in most evaluation systems. But the overriding question is how do we attach a 'value' or determine the 'worth' of what a teacher does in the classroom? Most teachers argue that the merit and worth of their contribution cannot be

seen immediately and neither can it be measured based on student's academic achievement in school.

In addition to the purposes given above, literature indicates other rationales for teacher evaluation, which include "instructional improvement, personal growth, accountability and control, promotion or dismissal and improvement of teaching performance" (Duke, 1995, p.4). Peterson (2000) gives various other reasons such as providing input for future teacher education programs, teacher reassurance, protecting children in school and to assure all those who are interested in teacher performance, in general.

According to Stake (1989), data gathered for teacher evaluation can also provide input for awarding merit and improving weaknesses, in helping to choose the most qualified teachers and retaining good teachers, in promoting ongoing professional education for teachers and lastly to provide a better understanding of the overall functioning of the school. Furthermore, through evaluation teachers will be able to reflect critically on the effectiveness of personal and professional practice (Smith, 2001).

Duke (1995, p.4) states that accountability, which has been the "buzzword in education since the early Seventies" has been raised as another reason for teacher evaluation. In this context, teachers are held accountable for student learning. Duke further explains that the reasons for this approach are to determine only competent teachers are allowed in the classroom and to ensure taxpayers that the money invested on education is proportionate to gains in student achievement. On occasions when such results are not available, the public will question the system used to evaluate teachers in the public schools. According to Darling-Hammond (1990), the model of accountability has been incorporated in most teacher evaluation systems and this calls for certain expectations regarding how teachers should perform their duties in school.

The main contention on the accountability approach is the issue of teachers being held accountable for assuring student achievement regardless of other factors that influence student learning. The need to see tangible improvements in student learning has pressed for more evaluation based on the accountability approach (Duke, 1995, Reeves, 2004). Frymier (1998) argues that teachers should not be held accountable for student learning because learning is behavior and no one should be held accountable for other people's behavior. Students should be taught to be responsible for their own learning. This view is supported by the idea that one single teacher cannot be responsible for the success of a student in school (Duke, 1995). It is usually a collaborative effort of all teachers who have taught the student throughout the year.

According to Reeves (2000) strategic leadership must promote accountability without instilling fear in teachers. Administrators need to introduce indicators of accountability to ensure that teachers are aware of their roles and take the necessary steps in ensuring better student outcomes. School leaders cannot be result-driven if they want teachers to teach well in the classroom. They should not put accountability pressure on teachers. On the contrary, they must ensure that teachers are given guidance on how to improve student achievement because improving student learning is a collaborative effort in school.

The more acceptable rationale for teacher evaluation emphasizes the developmental approach (Glickman, Gordon & Gordon, 2004). The initiative to carry out teacher evaluation for the purpose of professional development became more popular in the 1980s and early 1990s (Duke, 1995). Elliot (1989) states that research carried out in the United Kingdom asserts that evaluation of teachers should be mainly for the purpose of professional development. He explains further that professional development should help one acquire the necessary skills to perform professional tasks, possess the right beliefs and attitudes, as well as reflexive powers.

The developmental approach takes account of the teacher as a professional within the education system and provides opportunities to grow professionally. Findings of evaluation based on this approach can help plan on-the-job learning, efforts to enhance existing skills and formal education where necessary. However, research indicates that the present system of evaluating teachers has very little effect on promoting professional development in schools (Loup, Garland, Ellet & Rugutt, 1996; Peterson, 2000; Stiggins & Duke, 1988).

Most schools implement teacher evaluation system that addresses both teacher development (which is a formative process) and accountability (which is summative in nature). Therefore, these two purposes usually coexist and sometimes create confusion and role conflict (Peterson, 2000). School administrators must ensure that information collected for one purpose is not unethically used for the other purpose (Stake, 1989). Duke (1995) argues that probably in theory both these purposes can coexist in the same system but in practice it will not be feasible.

Whatever the reason given for teacher evaluation, a review of the literature indicates that the current system is considered ineffective and it has very insignificant impact on student learning (Frase & Streshly, 1994; Peterson, 1995; Shinkfield & Stufflebeam, 1995; Stiggins & Duke, 1998). Teacher evaluation is seen as a function to serve the bureaucratic needs of the organization rather than improving instructional practice (Wood & Pohland, 1979).

How to Evaluate Teacher Performance

There are various methods of evaluating teachers in schools and it includes the administrator evaluation, peer review, self-evaluation, teacher competency tests, student ratings, student achievement tests, teacher portfolio, principal ratings, parent survey or complaints and documentation of teacher professionalism (Barrett, 1986; Dwyer &

Stufflebeam 1996; Peterson, 2000). Each of these methods informs only one aspect of the teacher evaluation process, and is therefore insufficient in isolation. In most cases not all the methods are used, but definitely a mixture of various types will provide more reliable data on teacher performance (Peterson, 2000).

Stakes (1989) agrees that a systematic evaluation of teachers can only be carried out using a variety of methods to collect data. This view is further supported by Darling-Hammond (1990, p.26) who says, "any single method is most useful in combination with others that complement the data it can provide, and that the set of methods likely to be most useful may vary from one purpose or function of evaluation to another". In current practice, administrator evaluation, which consists of one or two classroom observations and interview, seems to be the most common method of teacher evaluation (Peterson, 2004). Some of the common methods used to evaluate teacher performance in school will be discussed below.

Administrator Evaluation

School administrators carry out teacher evaluation regularly to monitor performance and provide feedback on teacher practice. Barret (1986) states that most of the time school administrators tend to use classroom observation of instructional practice to evaluate both experienced teachers and beginners. Furthermore, interviews are used together with observation to have face-to-face discussions during pre-observation sessions and to communicate evaluation results to teachers. According to Burgess (1989), interviewing and classroom observation can be considered important components of the evaluation process though there are other aspects to an evaluation process. But Barret explains that through classroom observation administrators are able to see teacher interaction and rapport with pupils, which might not be possible with other methods.

Classroom observation can include peer observation, and observation by a mentor. However, there are mixed views about the exclusive use of classroom observation for teacher evaluation because, literature indicates that it can be a biased, invalid and unreliable form of collecting data (Darling-Hammond, Wise, & Pease, 1983, Marshall, 2005). If an evaluation based on classroom observation is going to be seen as a fair judgment of teacher competence, then it has to be based on frequent visits to the classroom (Hunter, 1988a). Marshall (2005) agrees that frequent short visits to the classroom can benefit the teacher evaluation process and provide more authentic data about what happens in the classroom.

Stodolsky (1990) believes classroom observations are useful when it comes to studying teacher behaviors or action, but it cannot contribute much insight into how teachers think, plan their teaching and prepare their materials, collaborate with their colleagues and parents or attend to student's needs. Finally, as Holly (1989, p. 108) aptly puts it, "the narrowness of focus and expectation as to what should happen in the classroom narrows the possibilities of what could happen". Administrator evaluation based on classroom observation should be combined with other methods of evaluation to get a more accurate picture of how teachers perform their jobs (Marshall, 2005; Peterson, 2000).

School leaders are encouraged to carry out more classroom walkthrough so that they can collect additional data on how instructional practice enhances student learning. Classroom walkthrough is designed to look for specific indicators of practice and it can last a very short period or cover the whole lesson (Poston, 2004). This strategy is carried out frequently and it is considered as a professional development activity which can help teachers improve their pedagogical skills (Barnes & Miller, 2001; David, 2008).

Peer Review

Another method of obtaining data on teacher evaluation is through an on-going peer review process that involves all teachers in the school (Darling-Hammond, 1990). Darling-Hammond et al. (1983) explained that through peer review teachers are supposed to have control of both the conditions of teaching and the quality of instruction. They will be able to identify strengths and weaknesses in the classroom and share ideas on improving instructional practice (Darling-Hammond et al., 1983).

Barrett (1986) believes peer review could help examine a wider scope of teaching activities since colleagues will be able to provide more precise assessment about instructional practice and student learning. But Peterson (2004) states that the question of reliability of observation will arise since teachers themselves will be only able to make a single visit to the classroom due to time constraints, and therefore he suggests that peer review should only include review of teacher materials. This method of collecting data will definitely take up too much of the teachers' time and in cases where there are disagreements about the findings it will result in conflict among colleagues.

Self-Evaluation

This method allows them to have more control and participation in their own evaluation (Peterson, 2000). The practice of self-evaluation can help teachers analyze their teaching and identify the area in which specific help is needed. Self-evaluation is a systematic and self-critical learning process for the teachers, with its ultimate goal to promote professional development (Darling-Hammond et al., 1983). It encourages them to reflect upon their teaching and find areas they need to improve, take the appropriate steps and report what changes they have carried out in their classroom.

Barrett (1986) explains that self-evaluation is usually used together with other more formal methods to detect deficiencies in instructional practice and classroom

management skills. Elliot (1989) believes that self-evaluation can play a key role in developing professional competence. This method provides teachers with an opportunity to express opinions on how they are improving their practice in the classroom and generate a sense of ownership over the process of teacher evaluation generally. When external parties evaluate and offer constructive criticism, they might not react positively towards it. That is why usually evaluation by external parties is often contrasted with self-evaluation (Elliot, 1989).

Self-evaluation activities can provide "the necessary psychological frame for ongoing self-improvement as well as the information needed to perceive events accurately and design avenues for change" (Darling-Hammond, 1990 p.27). In addition, self-evaluation, if it is reflective and emancipatory, will engage teachers and school administrators in an attempt to create a partnership for the improvement of instructional practice and the overall enhancement of the school. Darling-Hammond (1990, p.27) notes that the inclusion of self-evaluation in formative evaluation can "awaken and launch the most powerful source of improvement: the teacher's own capacity of self-evaluation".

Elliot (1989) advocates the same idea when he says that self-evaluation is an essential aspect of the process of personal development. Similarly, Airasian and Gullickson (1997) believe that self-evaluation is an integral part of an overall evaluation of professional growth and development. However, Darling-Hammond et al., (1983) cautions that though self-evaluation provides essential input for staff development, it cannot be used for accountability decisions. Administrators must use performance data from a number of sources to determine teacher accountability in school.

Teacher Competency Test

Teacher competency test is another method of teacher evaluation, which is widely used in the United States and other countries. The National Teachers Examination (NTE), which is used for initial certification and hiring decisions, is an example of competency-testing (Barrett, 1986). Advocates of using teacher tests argue that this method assures a basic knowledge level, removes interviewer bias, and is legally defensible (Darling-Hammond et al., 1983). However, the weakness of this method lies in its lack of validity and low correlation of the test results to teacher performance. Furthermore, the intangible aspects of teacher performance are rarely measured in these tests (Barrett, 1986). Therefore, this method is rarely advocated in the practice of formative evaluation.

Lately the teacher competency test has been introduced as a form of summative evaluation for teachers in Malaysian schools and it has created a lot of dissatisfaction among teachers. The test is administered by external evaluators from the Ministry of Education and it is conducted to select suitable teachers for promotion to a higher level and for increment in salary. The point of contention in this method of evaluation is that the test does not measure a teacher's actual competence. Most of the questions in the test are not related to what the teacher does in the classroom. The questions include financial matters, administrative circulars, and various other duties performed in school. This has lowered teacher confidence in the test and its resultant outcome.

Student Ratings

The use of student rating to evaluate teacher performance has not been widely used except in higher education. This method does not involve high cost and it is reliable but the issue of validity and bias remain (Darling-Hammond et al., 1983). In a study conducted to examine the use of student surveys to obtain information on teacher

quality at the K-12 level, the researchers found that student responses were reliable and valid (Peterson, Wahlquist & Bone, 2000). Therefore, they believe that student reports can provide important feedback on teacher performance and effectiveness in the classroom.

However, this method has hardly been used to evaluate teachers in Malaysian schools. In the primary school context the students are perceived to be too young to assess their teacher's performance and the issue of favouritism will make the findings invalid. In order for the findings to be a fair judgment of the teacher's performance it should be carried out by the teachers themselves as part of the self evaluation process and the must ensure that the student's identity is not revealed.

Student Achievement Tests

In many countries standardized national examination results have been used to assess teacher performance (Barrett, 1986). Student achievement is used as an indicator of teacher performance in occasions when there is a need to supplement data on outcomes of teaching (Glass, 1990). Peterson (2000) explains that student achievement can only be used in teacher evaluation if the data are defensible and obtained based on specific conditions such as agreement on how to measure student achievement, permission from teachers to use such data and the use of valid tests.

Research indicates that only under certain conditions, is there a positive correlation between test results and teacher behavior (Woolever, 1985). Many believe that student achievement depends on other factors besides teacher influence. Therefore, using achievement scores to evaluate teacher performance might raise issues of validity (Darling-Hammond et al., 1983). This point is supported by Eraut (1989) who explains that students come to the classroom with different capabilities and this will result in diverse learning outcomes. Thus, teachers should not be evaluated based upon the

learning outcomes of their students, "for teachers like other professionals, do not have control over all the variables that affect the outcome" (Smith, 1987, p.13). Therefore, in certain cases student achievement is not considered a good criterion to evaluate teachers (Barrett, 1986).

In most Malaysian schools student achievement in national examination is used to evaluate the performance of the school and indirectly the performance of the teachers. When the achievement level declines administrators and teachers are held accountable for it. They have to explain why the students did not perform well and take action so that it does not occur in the future. The examination oriented school system in Malaysia has increased the impact of student achievement in school. But teachers are not directly evaluated based on student achievement in the school-based teacher evaluation system.

Teacher Portfolios

A teacher portfolio consists of a teacher's own collection of information and material about his or her practice (Bird, 1990). It is believed that portfolios can present a comprehensive view of teacher practice in the classroom (Wolf, 1991). Research indicates that portfolios can be used to gather data about teacher performance in the areas of assessment and professionalism but the feasibility of using portfolios is a matter for concern due to the demands on teacher time in preparing the portfolio (Tucker, Stronge, Gareis, & Beers, 2003). Due to various problems encountered with evaluating teacher portfolios, it has been suggested that the teacher dossier would be a better option. The dossiers, which are more reliable and cost-effective, provide objective data that are easier to assess (Bird, 1990). Teachers in Malaysia are encouraged to have teacher portfolios but most often they do not update their portfolio because they are deluged with paperwork related to student learning and thus can hardly find the time to compile information for their individual portfolio.

Instruments and Criteria Used For Evaluation

Most of the teacher evaluation systems use multiple data sources in judging teacher performance. The various types of data available for evaluating teachers include classroom observation checklist, student achievement scores, peer review of materials, portfolios, teacher test scores, student and parent surveys and other documents (Peterson, 2000). The commonly used evaluation instrument appears to be the observation checklist and teacher rating forms. The instruments consist of various criteria or state mandated standards, which are used to judge teacher performance. Standards or criteria are needed to evaluate teacher competence and these criteria must be measurable (Barrett, 1986).

A review of the literature indicates that there is a lot of disagreement on the most appropriate criteria for teacher evaluation. The teaching methods of different subject teachers vary according to the subject matter and therefore it is not suitable to use the same generic criteria to evaluate all teaching situations (Barrett, 1986; Stake, 1989). Most researchers agree that the criteria used to evaluate teacher performance should include pedagogical and subject matter knowledge, teaching competencies, classroom management, teacher professionalism, ability to create a suitable classroom environment, good rapport with parents and students and other equally important duties such as contribution to the community and profession (Bridges, as cited in Dwyer & Stufflebeam, 1996). According to Burgess, (1989) we need to analyze critically all criteria used to evaluate teachers since there is no consensus on the best criteria that could possibly be used to assess teacher competency or performance.

The main concern in any effective teacher evaluation method is validity. The method, including the criteria or standards used, must measure what it purports to measure (Barrett, 1986). According to Eraut (1989), the evaluator must also know the subject matter; otherwise the validity of his judgment will be questioned. We must also refrain from making inferences from invalid measurements. The sparse research

concerning the validity of evaluation instruments and current practices used by most schools to measure teacher performance has created wide speculation and varying perceptions about the validity of performance instruments as well as professionals doing the evaluation. Most researchers insist on reliability and validity of instrument and criteria used to judge teacher performance. While validity and reliability are problems for personnel evaluations in general, they are accentuated in areas such as teaching because of the lack of consensus on the most appropriate criteria to use to assess teacher performance.

Utilization of Evaluation Findings

Stake (1989, p.29) believes that evaluation findings are usually used for "political manipulation" and hardly ever used to rectify mistakes in the system. Peterson (2000) supports this view when he states that administrators file up their reports and the public don't get to see them. Little effort is made to disseminate information about best teaching practice, special contributions by the teachers and generally success attained in the work place. This is contrary to the belief that teacher evaluation results should be utilized to promote professional development and to reassure the public that children in school are getting quality education from professional educators who are qualified and dedicated to their profession (Peterson, 2000). Evaluation findings can be used as input for staff development programs in school (Holland, 2005). It can be used to promote teacher growth and improve instructional practice in school. How administrators use the findings will determine the effects of teacher evaluation on teacher performance and eventually student learning in school.

Administrator's Role in Teacher Evaluation

School administrators are responsible for the leadership and management functions in their organizations. They should establish a shared vision and create a sense of direction for the school community to move forward in one direction to achieve common goals that ensure organizational success (Leithwood, Aitken, Jantzi, 2006). Administrators play a key role in leading, supporting, supervising and coaching as well as coordinating teaching and learning in their respective schools (Blasé & Kirby, 2000; Glickman, Gordon & Gordon, 2004). The main purpose of these functions is to enhance teacher performance and improve student learning in schools (Holland, 2005; Hunter, 1988b).

Research indicates that school administrators can positively influence teacher performance in school (Blasé & Kirby, 2000; Marshall, 2005). This can be done through collaborative actions, peer coaching, reflective dialogues and classroom research (Blasé & Blasé, 1999). They can also promote organizational learning and improve school effectiveness if they plan and implement effective supervision and evaluation strategies in schools (Marshall, 2005).

Principals and head teachers are entrusted with the responsibility of being instructional leaders who develop curriculum, supervise classroom teaching, monitor student achievement, plan staff development and generally assist teachers in implementing the curriculum in the classroom (Blasé & Blasé, 1999; Glatthorn, 2000). Blasé and Kirby (2000) explain that principal's action can affect instructional climate and subsequently this can have an effect on student achievement outcomes in schools. Literature on instructional leadership in schools indicates that there is a strong link between leadership behaviours, teacher commitment and professional involvement of teachers in school (Blasé & Blasé, 1999; Blasé and Kirby, 2000; Glatthorn, 2000). School administrators as curriculum leaders must also provide adequate resources,

observe the implementation of the curriculum, take appropriate measures to improve student test scores and plan teacher development program (Glatthorn, 2000). How administrators communicate with teachers and provide positive feedback will also have an effect on teacher practice in the classroom. Furthermore, regular praise and proper modeling will help develop teacher professionalism in school. Research also indicates that transformational leadership can improve teacher behaviour, attitude and effectiveness in schools (Glatthorn, 2000; Leithwood, Aitken & Jantzi, 2006; Reeves, 2002). Therefore the different leadership roles played by school administrators can have a profound effect on teacher quality and student learning.

The diverse roles played by the school leaders have created a situation where they are responsible for both the instructional supervision and evaluation of teachers in school (Glanz & Sullivan, 2000; Glickman et al., 2004; Holland, 2004). On the one hand, they are supervisors who are supposed to help teachers improve their instructional practice, and on the other, they are evaluators who must make value judgments on teacher performance. In such a situation how would teachers reveal their problems to a 'supervisor' and conceal it from an 'evaluator' (Hunter, 1988b)? This creates a conflict in the roles played by these administrators (Glickman et al., 2004, Holifield & Cline, 1997). Regardless of this, they still have to establish and maintain personnel management practices that include evaluating the performance of teaching staff and providing appropriate feedback to improve practice.

Challenges in Teacher Evaluation

Research indicates that there are numerous problems in the implementation of teacher evaluation system in schools (Holland, 2005; Marshall, 2005; Peterson, 2004). The ineffectiveness of the teacher evaluation system seems to be the main problem in most countries (Holland, 2005; Marshall, 2005; Odhiambo, 2005; Peterson, 2000). This

could be associated with insufficient resources, bureaucratic practices, inadequate training for evaluators, lack of suitable instruments, lack of commitment from administrators and very little information from validation studies (Dwyer & Stufflebeam, 1996; Holland, 2005; Marshall, 2005).

Peterson (2000) states that 70 years of research in the United States did not indicate any empirical evidence to prove that teacher evaluation has brought about improvement in instructional practice in the classroom. He strongly believes that the current system does not improve teacher confidence, identify innovative teaching or provide any form of reassurance about teacher quality.

According to Burgess (1989, p.25) studies done on teacher performance by psychologists indicate that there are several problems related with defining and measuring teacher performance. Another problem which requires mention seems to be the gender issue. Burgess (1989) has raised concerns about gender issues in teacher evaluation. He explains that most teachers are women, whereas the evaluators are men and this might give rise to a male-dominated point of view, which might be considered bias by some. The other concern in teacher evaluation centers on who conducts the evaluation, the criteria used and how the information is collected. The problems are also related to several other aspects such as principal rating, classroom observation, interviews, content of assessments, and evaluation imposed from outside.

When school principals or administrators are evaluators, then the hierarchical relationship that exists in school will give rise to the question of power and authority in gathering information pertaining to teacher performance and how this information will be used (Burgess, 1989). Research indicates that principals' reports are still questionable (Marshall, 2005; Peterson, 2004). Studies conducted from 1921-1959 showed that there was very low statistical correlation between principal ratings of teacher performance and teacher effectiveness (Medley & Coker, 1987). Most administrators consider teacher

evaluation the least popular part of their duties and are therefore not really motivated to do a good job (Peterson, 2000). Generally, principals are unable to evaluate accurately because they cannot determine what constitutes good teaching and even the checklist form they use is based on assumptions about good teaching (Peterson, 2000).

The most common method used by administrators is the classroom observation where an unreliable checklist is used to evaluate how teachers carry out their lessons (Peterson, 2004). Consequently, the question that arises is, whether they are able to make accurate judgments about teacher performance and support it with reliable data gathered using this method. Teacher evaluation based on brief classroom observation cannot paint an accurate picture of teacher performance (Marshall, 2005; Medley & Coker, 1987; Scriven, 1981; Stodolsky, 1990). This could be due to the fact that the administrators themselves are not experts in the subject matter content and evaluator judgment depends heavily on what the teacher is doing at the moment (Peterson, 2000).

Scriven (2003, p.303) states that "direct observation alters the behavior of the observed teacher" and thus, it cannot be considered as valid evaluation of what the teacher does in the classroom. Most administrators have very little time to conduct multiple observations due to their pressing need to attend to managerial duties. Therefore they are unable to verify if a teacher's performance is consistent in the classroom and this, results in unreliable data for evaluation (Marshall, 2005; Peterson, 2000).

Only during certain clinical supervision is the complete process of pre-observation conference, discussion on elements to observe, observation and post-conference conducted. Hunter (1988b) also says that many administrators are unable to evaluate or supervise teachers properly. She believes that this could be due to lack of training as evaluators. Furthermore, criticism at performance-based teacher evaluation points out that the classroom context is not taken into account during assessment and this

makes the evaluation system ineffective (Darling-Hammond et al., 1983; Marshall, 2005; Stodolsky, 1988).

Another problem with teacher evaluation is that interviews carried out with teachers might not be able to provide valid information due to the hierarchical relationship between the teacher and the evaluator and the "status of confidentiality of the interview" (Burgess, 1989). The teachers are worried about how the information will be used in the future. Thus, they only give positive responses, which they perceive the evaluator would want to hear.

There is no agreement on what should be measured when we evaluate teacher performance because the views on what is a suitable knowledge base for teaching given so far do not help in constructing suitable instruments for evaluating teacher performance (Dwyer & Stufflebeam, 1996). In order to evaluate the quality of teacher performance, evaluators have to possess some knowledge of teaching (Stake, 1989).

Literature indicates that professional standards based on sound professional practices of evaluation have been designed to assess and improve teacher performance (Dwyer & Stufflebeam, 1996). Standards must be validated and also periodically reviewed and updated. In evaluating teachers we must also ensure that they don't just have competency or certain appropriate behaviors, but they must behave appropriately to suit the context in the classroom (Peterson, 2000).

Teacher Evaluation Models

According to Dwyer and Stufflebeam (1996), there are various models of teacher evaluation based on different rationales, which include "improvement of classroom teaching; professional accountability and development; administrative supervision; research–based indicators for improving student outcomes; consumer protection and public responsiveness; and merit pay" (p.774).

The improvement model emphasizes regular observation of instructional practice with constructive criticism as feedback to improve practice. Administrators observe and record data on teacher pedagogical skills, teaching competences, subject matter knowledge and classroom management. The outcomes of this observation will result in regular professional development programs.

Models based on professional accountability emphasize the need for continuous improvement in the service provided to students and encourage teachers to present supporting data pertaining to the quality of their service (Dwyer & Stufflebeam, 1996). The administrative supervision model requires principals to evaluate teacher performance to make sure that teachers have carried out their responsibilities and taken steps to strengthen their performance.

On the other hand, the research-based indicators model focuses on assessing teaching behaviors that correlate positively with student achievement. According to Dwyer and Stufflebeam (1996) many state teacher evaluation systems have adapted this model and used measurement-based teacher performance in their evaluation systems.

The consumer protection and public responsiveness model promotes teaching as a public service. Thus, effective teaching and student welfare is given prominence and high standards are set for teacher performance (Dwyer & Stufflebeam, 1996). The model which advocates better pay for good performance is called the merit pay model. In this model teachers are evaluated based on student achievement and supervisor's judgments.

Seyfarth (1991) gives an alternative explanation to the various models of teacher evaluation discussed above. He explains that the three models that are used extensively in schools include the remediation model, the goal-setting model and the product model. According to Seyfarth (1991), the remediation model offers planned remediation sessions for teachers who did not achieve high ratings in their summative evaluation. In this model administrators evaluate, provide feedback and reevaluate teachers to help

correct observed weaknesses. He cautions that this model requires more time and could be a burden for the evaluators.

In the goal-setting model, teachers and administrators are involved in determining the evaluation criteria. The teachers are given the opportunity to set the developmental goals and strategies for achieving them. The administrators review, make changes and endorse it before finally evaluating the attainment of the goals. The challenge faced in implementing this model is in setting appropriate goals that meet the expectations of the administrators. This model gives more autonomy to teachers and usually works well for formative evaluation (Seyfarth, 1991). Finally, in the product or outcome model student achievement is used as indicator to evaluate teacher performance. Teachers are held accountable for student learning and therefore student achievement in pre- and post-tests are taken into account to evaluate teacher effectiveness in the classroom. This model is acceptable if student achievement is measured using valid instruments and teachers teaching weak students are not penalized for low student achievement (Seyfarth, 1991).

The use of the most appropriate teacher evaluation model depends on the purpose of teacher evaluation within the school system. The investigation in this study included the search for a Malaysian teacher evaluation model that was prevalent in the national primary school. It was important to ascertain current practices in order to introduce a suitable model for the local context.

Overview of Teacher Evaluation Practices in Malaysia

The education system in Malaysia is centralized and most policy decisions are made by the Federal government (Ministry of Education, 1985). The Ministry of Education has the responsibility of developing a national education system to provide quality education for all children for a minimum of 12 years. This national system has been carefully designed to provide a common curriculum, standardized national

examinations, and the use of the national language as the main medium of instruction (Ministry of Education, 2001).

Primary and secondary education is highly structured, with the purpose of providing all students equal opportunity to a solid basic education and to equip them with fundamental skills that will enable them to adapt to the changing world outside school. To achieve this purpose the Ministry of Education has equipped schools with well-trained teachers who possess the necessary competence to educate all children to become vital national resource for the development of the nation (Ministry of Education, 1997). The Ministry of Education has given prominence to pre-service and in-service teacher training in order to develop a corps of professional teaching personnel who will provide quality education in schools.

Teacher evaluation is considered an essential function to judge teacher competence in carrying out their professional duties and to provide feedback for further improvement of instructional practice in schools. However, the history of teacher evaluation practices in Malaysia reveals that though the practice of assessing teachers has been conducted for some time, there is no clear indication of how the formal school-based teacher evaluation system was defined and carried out, or how it contributed towards improved instructional practice in Malaysian schools.

In the early days, the teacher evaluation system was not as well-developed as in the United States or Britain. Teacher evaluation was considered part of the functions of the Federal Inspectorate of Schools, which advocated supervision and inspection as a means of improving instructional practice in schools (Chan, 1994). In an effort to ensure that high academic and professional standards were maintained among the teachers, the Ministry of Education established a system to inspect, monitor and supervise teachers regularly. In 1956, the Federal Inspectorate of Schools, a separate independent department which is directly responsible to the Minister of Education, was established to

conduct school inspection. Its function included observing, monitoring and advising teachers to ensure that the quality of teaching and learning, as well as the administration in schools were maintained and improved (Ministry of Education, 1985).

The Federal Inspectorate of Schools conducts full inspection, follow-up inspection, thematic inspection and special inspection to assess all aspects of school, which includes evaluating teacher performance. However, owing to the increase in the number of schools in the country and a shortage of school inspectors, the responsibility of evaluating teachers became the prerogative of the respective school administrators. Headmasters and senior assistants, by virtue of being school administrators, conducted teacher evaluation in their schools. This started the decentralization of some of the functions of evaluation from the Federal Inspectorate of Schools to the school administrators who used the concept of supervision to form the basis for their school-based evaluation system.

The term teacher evaluation is hardly ever used directly in Malaysia. Terms such as 'classroom observation', 'monitoring' and 'supervision' are used instead. School administrators conduct evaluation based on guidelines given in departmental circulars and use their own instrument or instruments adapted from other agencies to gather data on teacher performance. According to Chan (1994), the evaluation system evolved over the years and there were several home-produced schemes which were developed by headmasters who had attended management training programs.

Formative evaluation is carried out to determine the effectiveness of instructional practice in the classroom and the overall quality of the teacher's work performance. Besides this, summative evaluation in the form of teacher appraisal is conducted at the end of the year. Administrator ratings are used to evaluate teacher performance and the findings are used to make decisions on promotion, offer financial inducement and better opportunities to upgrade professionally.

The introduction of the New Primary School Curriculum in the 1980s brought about the full implementation of the teacher evaluation system in most Malaysian primary schools (Chan, 1994). Some school administrators were provided training in management by the National Management Training Institute and components of the course included teacher evaluation and supervision. The school-based evaluation system, which compliments the functions of the Federal Inspectorate of Schools, has helped attain the long-term objective of providing quality education for all children (Chan, 1994).

Latest development in teacher evaluation in Malaysian schools indicates that whenever a teacher has been proposed to be conferred the 'Master Teacher' post, the evaluation of the teacher's performance is carried out by the Federal Inspectorate of Schools and not by school administrators. This is because the results of the evaluation can affect decisions on promotion and merit pay.

Other changes include the introduction of the Malaysian School Education Quality Standard (*Standard Kualiti Pendidikan Malaysia-Sekolah)* by the Federal Inspectorate of Schools to encourage schools to conduct self-assessment and take necessary steps towards overall school improvement (*Kementerian Pendidikan Malaysia,* 2003*).* This standard has been introduced in schools since 2003 and has been revised over the years. Most schools use the instrument for the teaching and learning dimension of this standard as a benchmark to evaluate teaching and learning in their schools. Further inquiry has to be carried out to investigate if this new self-assessment process for the entire school has brought about any changes in the evaluation of teachers and to what extent teachers are supportive of this new system. Unfortunately, very little research has been carried out to look in–depth at this issue, especially from the teachers' perspective.

In Malaysia, teachers are employed after completing their teacher education courses and placed in schools, and they usually remain in service until they choose to

quit or when it is time for compulsory retirement. They are hardly ever removed or dismissed from their position based on evaluation results. So the question of whether evaluation results should be used to determine tenure, dismiss incompetent teachers or retain good ones hardly ever arises in the Malaysian context.

Research indicates that this situation arises because the teacher evaluation system has been developed and implemented for the purpose of public accountability, and in some instances the system is seen to be conducted to meet administrative needs and not to upgrade teacher professionalism (Chan, 1994). There is a need to include the aspect of professional development in the evaluation system to enhance personal performance and improve instructional practice (Duke & Stiggins, 1990). It has been found that developmentally oriented evaluation not only impacts instructional practice, but can also improve teacher motivation and job satisfaction (Duke & Stiggins, 1990; Duke, 1995; Dwyer & Stufflebeam, 1996; Natriello, 1990; Peterson, 2000).

Furthermore, debates and discussions pertaining to evaluating teachers have brought to the fore various challenges in the implementation of the teacher evaluation system in Malaysian primary schools. These challenges include the lack of training for evaluators; sole use of classroom observation to evaluate teachers; absence of self evaluation; use of unreliable instruments; suitability of the criteria used for evaluating teachers; lack of various data sources and superficial evaluation due to heavy workload of school administrators (Abdul Aziz, 1990; Chan, 1994; Chan, 1997). This study examined how these challenges have been addressed over the years in the Malaysian national primary schools.

Review of Studies on Teacher Evaluation

There is extensive empirical literature on teacher evaluation and the appraisal of educators from the west, especially the United States and the United Kingdom. The

findings of these studies are also relevant to researchers in Malaysia in order to understand further the practice of evaluating teachers and using the results to improve instructional practice. This section will review a few pertinent studies, which range from major international studies comparing practices in different countries to some small studies carried out by individuals. The review of teacher evaluation brought forth various debates on issues pertaining to the purpose of teacher evaluation, methods or procedures, role of evaluators, criteria for evaluating, instrument used, impact on instructional practice and the overall effectiveness of teacher evaluation. It also highlighted problems in practice that need to be rectified so that teacher evaluation can play an important role in the professional development of teachers.

An international study involving five countries (Cyprus, Hong Kong, The Netherlands, UK and the USA) was carried out to compare relationships among educational effectiveness, evaluation and improvement variables. The result of this study indicated that three distinct purposes for conducting teacher evaluation in these countries were accountability, promotion and staff development (Teddlie, Springfield & Burdett, 2003). This study also revealed that more research needs to be conducted in these countries to investigate the relationship between teacher evaluation and other constructs, such as staff development and educational improvement.

Several other studies looked at the various methods used in teacher evaluation. A study conducted in Hong Kong to gauge teachers' and principals' perceptions of using classroom observation for staff development and evaluation noted that it was mainly for evaluation rather than staff development. The survey results reported that elementary teachers were not too keen to have observers in their classroom (Lam, 2001).

A review of the literature also indicates that alternative methods of teacher evaluation have resulted in more positive reactions from teachers. The findings of a qualitative case study of a horizontal teacher evaluation program in an urban school

district supported the claim that formative evaluation can result in changes in teachers (Johnson, 1998). In this study, teachers were involved in observing one another and having in-depth conversation or dialogues about teaching aims and practices. All teachers who participated in this study demonstrated that they could improve their classroom practice through discussion with peers and furthermore, they also overcame negative attitudes towards teacher evaluation (Johnson, 1998).

Another method that has been suggested for gathering teacher evaluation data is the use of student survey or student rating. Research by Peterson et al., (2000) was carried out to test a student survey instrument which could be used for teacher evaluation. This large-scale empirical study examined individual items in the questionnaire to determine norms to interpret future use of student views. An item analysis was carried out and it revealed that students responded to the items with reliability and validity; thus, it was concluded that student surveys could be reliable data source for teacher evaluation (Peterson et al., 2000).

A case study on instructional supervision, including teacher evaluation, was carried out in an urban, high performing high school. The findings of the study showed that formal evaluation had little effect in improving instructional practice, but respondents indicated that they needed critical, instructional feedback (Astor, 2005). Whereas, concurrent studies carried out by a group of doctoral students on teacher evaluation and its effects on instructional practice, produced varying results. These studies used the same research questions, methodology and instrument for research, but the sites of research differed. A case study in a high performing urban elementary school noted some effectiveness in the teacher evaluation system practiced in this school but other factors such as teacher efficacy, teacher collaboration with peers and effective hiring practices had a significant effect on teacher practice (Muhonen-Hernandez, 2005).

Findings of a study conducted in an elementary school indicated that the teacher evaluation process was being fully implemented and the evaluation process could play a better role in improving teacher practice if there was increased collaboration, informal administrator observation, and special training for evaluators (La Masa, 2005). Another case study in a high performing, urban elementary school concluded that though the teacher evaluation tool used in this school was effective, the evaluation itself had little impact on the overall improvement of the school. Furthermore, it indicated that four aspects that were perceived to have a greater impact on student achievement was staff collaboration, quality leadership, professional development and an emphasis on student achievement (Hillyer, 2005).

Another study on teacher evaluation and its impact on teacher practice in an urban high performing elementary school found that the evaluation process in this school supported teacher practice, and self reflection by teachers helped further improve teacher practice (Santos-Camerino, 2005). The study indicated several other factors, such as, strong instructional leadership, effective professional development, peer collaboration, focus on student achievement, decisions based on data and parent involvement had strong impact on teacher practice (Santos-Camerino, 2005).

A look at the use of standards in teacher evaluation denotes that there are various types of standards used and some are more effective than others. A study was conducted to determine whether teachers using state-mandated teacher evaluation systems or locally developed alternative teacher evaluation systems perceived teacher evaluation as having a stronger positive impact on school improvement, professional development, and student learning. This study used both the qualitative and quantitative methods to collect data using the Teacher Evaluation Profile (TEP) from 4,092 teachers in 21 school districts. Qualitative data were gathered from interviews with the personnel director in each district. Locally developed alternative teacher evaluation systems were perceived as

having stronger impact on school improvement, professional development, and student learning than state-mandated evaluation systems (Colby et al., 2002). Findings of the study also indicated that locally developed evaluation systems were better able to support district reform initiatives, guide professional development, and use student learning as a focus for teacher evaluation.

Denner, Salzman and Bangert (2002) carried out a study to examine the link between teacher assessment and student performance. They studied the validity and generalizability of using 132 teacher work samples to evaluate the ability of pre-service teachers and in-service teachers to meet national and state teaching standards and to make an impact on student performance. The findings of the study indicated that teacher work sample assessment could be used to give valid and credible evidence to show that teacher performance can affect student learning.

A review of studies conducted in Malaysia showed that research on teacher evaluation in Malaysia either looked at formative evaluation, which consists mainly of classroom observation or summative evaluation, which concentrated on appraisal of teachers for personnel decision-making (Abdul Aziz, 1990; Chan, 1994; Chan, 1997; Mohd Zakaria, 2002). It was also clear that most of these studies have been carried out based on individual interest in the area rather than a concerted effort or policy by the government. Due to the sparseness of available research only tentative conclusions can be made about the effectiveness of teacher evaluation practices in Malaysia. Generally, teachers and other educators were dissatisfied with the evaluation practices, and felt that existing criteria and methods used to judge teacher performance needed modification and improvement (Abdul Aziz, 1990; Chan, 1994; Chan, 1997; Mohd Zakaria, 2002).

The teachers were not really against the idea of being evaluated, but they had a sense of dissatisfaction with the way it was conducted (Chan, 1994). The researchers also felt that there should be more teacher involvement in the process and transparency in

using the results of the evaluation. The findings of these studies indicate that though teachers felt that classroom observation should not be the only method for formative evaluation, they felt that classroom visitation reports should be included in the summative evaluation which is carried out at the end of the year (Abdul Aziz, 1990; Mohd Zakaria, 2002).

The existing body of research on teacher evaluation indicates that various factors such as, training provided to the evaluators; standards and criteria used; resource and time allocated for evaluation; professional development planned for teachers and teacher cooperation in the overall process, contribute towards the success of teacher evaluation practices in schools (Drake, 1984, as cited in Barrett, 1986). There is also a need to understand what constitutes good teaching quality in education and the personal qualities that a teacher must acquire to teach well (Elliot, 1989).

A review of the literature further reveals that teacher evaluation is important and we need to conduct it for the purpose of professional development. Formative evaluation if carried out effectively will not only reassure the public of teacher effectiveness but also motivate teachers to improve their performance. There is still a lack of information on best practices in teacher evaluation and much work remains to be done. It is time to investigate how to evaluate teacher performance appropriately according to the environment they work in and develop suitable models of evaluation to suit the context in respective countries.

Chapter Summary

This chapter reviewed the literature in the area of teacher evaluation. It discussed the main concepts of the study and the findings of other researchers in the area of teacher evaluation. The discussion covered areas such as instructional practice, the need for teacher evaluation, formative and summative evaluation, the rationale for evaluating

teachers, teacher evaluation methods, instruments and criteria used for evaluation, as well as utilization of evaluation findings. This was followed by a discussion of the administrator's role in teacher evaluation, challenges faced in implementing teacher evaluation practices, models of teacher evaluation and an overview of teacher evaluation practices in Malaysia. This chapter concluded with a review of empirical studies on teacher evaluation practices.