



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN AND

MATHEMATICAL SCIENCES

Southampton Education School



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

**An Evaluation of The Implementation of
The School-Based Assessment System in Malaysia**

by

Nor Hasnida Che Md Ghazali

Thesis for the degree of Doctor of Philosophy



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

UNIVERSITY OF SOUTHAMPTON

ABSTRACT



05-4506832



pustaka.upsi.edu.my



Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Southampton Education School

Thesis for the degree of Doctor of Philosophy

AN EVALUATION OF THE IMPLEMENTATION OF THE SCHOOL-BASED ASSESSMENT SYSTEM IN MALAYSIA

Nor Hasnida Che Md Ghazali

The implementation of the school-based assessment (SBA) system is an effort in improving human capital development in a holistic manner and also to lessen the negative influences of exam-oriented education systems on students. The need to evaluate the new system is of critical importance as the SBA system is still in a relatively early stage of development. The research is conducted to evaluate the implementation of the SBA system in Malaysian schools by using Stufflebeam's CIPP (context-input-process-product) Model. According to Stufflebeam, any programme could be evaluated from four dimensions of context, input, process and product. Applying the CIPP Model and supported by relevant learning theories - behaviourism, Piaget's learning theory, constructivism, multiple intelligence and brain research and the assessment models - formative model, the logic model and the SCAP (Social Constructivist Assessment Process) Model, this research examines the evaluation instrument in order to choose valid, just and quality items. The research also investigates the interrelationship of all the evaluation dimensions in the context of SBA implementation. Although much has been done to investigate the relationship between dimensions in this context, none has related all the dimensions together. Additionally, this research incorporated different types of school (urban-rural) and school category (secondary-primary) as variables, which possibly moderated the relationship between the evaluation dimensions. A stratified random sampling technique was applied to collect data from 776 teachers in primary and secondary schools in Kelantan, one of the states in the north-east of Peninsular Malaysia. All the dimensions of evaluation were measured using a questionnaire developed by the researcher. A structural equation modelling software called AMOS (Analysis of Moment Structures) was applied to test all the hypotheses of the study. Results of the study showed i) input dimension has met the desired result, process dimension could still be strengthened and production dimension showed that participants' needs were partially met; ii) evidence of measurement models for input, process and product. Results failed to reject the hypothesised relationships between input and process dimensions and also between process and product dimensions; iii) input had a significant relationship with process, some process components had a significant relationship with product; and there was an indirect relationship between input and product; and iv) school category was found to moderate the relationships between dimensions. Theoretical, methodological and practical implications are discussed. In short,



this study provides support for the effectiveness of SBA implementation in schools.



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

Table of Contents

ABSTRACT.....	i
Table of Contents	i
List of tables	vii
List of figures	xi
List of accompanying materials.....	xiii
DECLARATION OF AUTHORSHIP	xv
Acknowledgements.....	xvii
Definitions and Abbreviations	xix
Chapter 1: Introduction.....	1
1.1 Background to the study.....	1
1.2 Statement of the problem.....	3
1.3 The conceptual framework	6
1.4 Making contribution to knowledge	13
1.5 Purpose of the study	15
1.6 Research questions	16
1.7 Research hypotheses	17
1.8 Summary.....	18
Chapter 2: The Assessment System in the Malaysian Educational System.....	21
2.1 Introduction	21
2.2 The Change in the Assessment System.....	23
2.3 National Education Assessment System	26
2.3.1 Central examination.....	27
2.3.2 School-based Assessment	27
2.3.3 Central Assessment	29
2.3.4 School Assessment	30
2.3.5 Physical Activity, Sports and Co-Curriculum Assessment.....	34
2.3.6 Psychometric Assessment.....	34
2.4 SBA Training Plan.....	36
2.5 Quality Assurance	37
2.6 Management System of SBA.....	39
2.7 Latest SBA Implementation.....	39
2.8 Summary.....	40

Chapter 3: Literature Review	41
3.1 Introduction	41
3.2 Assessment	41
3.2.1 Type of Assessment	46
a) Formative Assessment	47
b) Summative Assessment	52
3.3 School Based Assessment	54
3.3.1 Definitions and concepts	55
3.3.2 Role of SBA	59
3.4 School Improvement	61
3.4.1 Definition of School Improvement	61
3.4.2 The Concept of School Improvement.....	63
3.5 Assessment Models	70
a) Formative Assessment Model	71
b) Logic Model	73
c) The Social Constructivist Assessment Process Model	76
3.6 Program Evaluation	81
3.6.1 Concept of Evaluation	81
3.6.2 Types of evaluation	84
3.6.3 Program evaluation and its purpose	86
3.6.4 Models of program evaluation	89
3.6.5 The CIPP Model	98
3.6.6 Reason for choosing the CIPP Model as a Framework of this study	105
3.7 Context Dimension of SBA Implementation Evaluation	107
3.8 Input Dimension of SBA Implementation Evaluation	108
3.9 Process Dimension of SBA Implementation Evaluation	109
3.9.1 Teacher's attitude towards SBA	109
3.9.2 Teacher's understanding on SBA	111
3.9.3 Effect of SBA courses in improving teachers' skills	113
3.9.4 In-house training on SBA	114
3.9.5 Encouragement by administration	115
3.9.6 Teachers' commitment towards Moderation process	115
3.9.7 Teachers' commitment towards Monitoring process	119
3.9.8 Challenges in SBA System Implementation	120
3.9.9 The Role of SBA System in School Improvement	125

3.10 Product Dimension of SBA Implementation Evaluation	128
3.11 Summary	130

Chapter 4: Research Methodology	131
--	------------

4.1 Introduction	131
4.2 Research Paradigms	131
4.3 Research Design	132
4.4 The Population and the sample	135
4.4.1 The Population of the Study	136
4.4.2 The sampling method	136
4.4.3 The sample of the schools	137
4.4.4 Teachers' samples	140
4.5 Methods of Data Collection	141
4.5.1 Instrument of the study	143
4.6 Validity and Reliability of the questionnaire	146
4.7 Pilot Testing	150
4.7.1 Findings of Pilot Study of the Questionnaire	151
4.8 Data Collection Procedures	160
4.9 Ethical Issues	161
4.10 Summary	163

Chapter 5: Analyses and Results	165
--	------------

5.1 Introduction	165
5.2 Overview of the data analysis	165
a) First stage: Assessing the Measurement Model	167
b) Second stage: Analysis the Structural Equation Modelling	168
5.3 Requirement for SEM analysis	169
5.4 Response Rate	170
5.5 Preliminary Data Analysis	170
5.5.1 Missing Data	171
5.5.2 Outliers	171
5.5.3 Assessing Normality	172
5.6 Profile of Respondents	174
5.7 Analyses and Results	176
5.8 Analyses and Results of Measurement Models: Stage 1	188
a) Estimation method	188

b) Goodness-of-fit Index Assessment	188
c) Uni-dimensionality	191
d) Validity	191
e) Reliability	192
5.8.1 Input Dimension	193
5.8.2 Process Dimension	200
5.8.3 Product Dimension	211
5.9 Analyses and Results of Structural Model: Stage 2	221
5.9.1 The Hypothesized Model - The Structural Model 1	222
5.9.2 The Hypothesized Model - The Structural Model 2	225
5.9.3 The Hypothesized Model - The Structural Model 3	225
5.9.4 The Hypothesized Model - The Structural Model 4	226
5.10 Moderating Effect of Observed Variables	230
5.10.1 Moderating Effect of School Type	230
5.10.2 Moderating Effect of School Category	234
5.11 Summary of Hypothesis Testing	237
5.12 Chapter Summary	239
Chapter 6: Discussion, Implications and Conclusion	241
6.1 Introduction	241
6.2 An Overview of the Study	241
6.3 Summary of the Findings	243
6.4 Discussion of the Findings	245
6.4.1 Input, process and product evaluation dimensions of the SBA system implementation	245
6.4.2 The measurement scale for input, process and product evaluation dimension	250
6.4.3 The relationship among evaluation dimensions	253
6.4.4 Moderating Role of School Type and School Category	261
6.5 Implications of the study	263
6.5.1 Theoretical Implications	263
6.5.2 Methodological Implications	265
6.5.3 Practical Implications	266
6.6 Limitations of the Study	270
6.7 Suggestions for Future Research	272
6.8 Conclusion	272

Appendix A: SBA Evaluation Questionnaire (English Version) 277

Appendix B: SBA Evaluation Questionnaire (Malay Version) 281

Appendix C: Consent Form (English Version)286

Appendix D: Consent Form (Malay Version)287

Appendix E: A Letter to Headteacher (English Version).....288

Appendix F: A Letter to Headteacher (Malay Version)289

Appendix G: Cover Letter (English Version)290

Appendix H: Cover Letter (Malay Version)291

**Appendix I: The Frequency Histograms and The Normality Probability Plots
.....292**

Appendix J: Missing data statistics by item315

Appendix K: SBA Evaluation Questionnaire for Pilot Study317

Glossary321

List of References322

List of tables

Table 1.1	Comparison of traditional and new concept of mathematics Assessment	5
Table 1.2	Types of Evaluation and Decisions	9
Table 1.3	The Framework of the CIPP Model	10
Table 1.4	The Context, Input, Process and Product Evaluation	12
Table 2.1	Public examination in the Malaysian education system	22
Table 2.2	Difference between KBSR and KSSR	24
Table 2.3	Content Standard and Learning Standard of KSSR	25
Table 2.4	SBA and public examination from preschool to upper secondary School	28
Table 2.5	Framework of the Performance Standard	32
Table 2.6	Framework of the Performance Standard of English Language Year 1	33
Table 2.7	Framework of the PSD	33
Table 2.8	PSD of Mathematics Year 1	33
Table 2.9	Four components of SBA	35
Table 2.10	SBA training plan	37
Table 2.11	Officers responsible for QA	38
Table 2.12	Work schedules of QA	38
Table 2.13	SBA and CE component for primary and secondary level	40
Table 3.1	Category of formative assessment	48
Table 3.2	Formative and summative assessment	53
Table 3.3	Summary of the learning theories with their implications	76
Table 3.4	The concepts and variables tested or applied	86
Table 3.5	The key strategies of formative assessment	88
Table 3.6	Main differences between formative and summative evaluation	90
Table 3.7	Educational programme	96
Table 3.8	Evaluation Approaches, Models and their Focus	102
Table 3.9	Level and its measurement focus in Kirkpatrick's model	103
Table 3.10	Formative and summative evaluations	103
Table 3.11	The relationship of evaluation roles to context, input, process and product evaluations	117

Table 3.12	Summary of the evaluation models	117
Table 3.13	The difference between external and internal moderation ...	119
Table 3.14	The implementation of the moderation process at the three stages	120
Table 3.15	The difference between external and internal monitoring	123
Table 4.1	The teacher-student ratio in Malaysia	135
Table 4.2	Number of schools, teachers and students	136
Table 4.3	The number of secondary and primary schools chosen	138
Table 4.4	Number of school, teacher and student in Kelantan	141
Table 4.5	Contents of the questionnaire	146
Table 4.6	Values of Cronbach's Alpha if item deleted and Overall Cronbach's Alpha for the Input Evaluation Constructs	152
Table 4.7	Two-factor pattern matrix showing the factor loadings of each of the variables	152
Table 4.8	Three-factor pattern matrix showing the factor loadings of each of the variables	153
Table 4.9	Values of Cronbach's Alpha if item deleted and Overall Cronbach's Alpha for the Process Evaluation Construct	154
Table 4.10	Three-factor pattern matrix showing the factor loadings of teachers' attitude construct	155
Table 4.11	Three-factor pattern matrix showing the factor loadings items.....	155
Table 4.12	Three-factor pattern matrix showing the factor loadings	156
Table 4.13	Two-factor pattern matrix showing the factor loadings of items	157
Table 4.14	Values of Cronbach's Alpha if item deleted and Overall Cronbach's Alpha for the School Improvement Evaluation Construct	157
Table 4.15	Two-factor pattern matrix showing the factor loadings of items	158
Table 4.16	Values of Cronbach's Alpha if item deleted and Overall Cronbach's Alpha for the Product Evaluation Constructs	158
Table 4.17	Three-factor pattern matrix showing the factor loadings	159
Table 4.18	All items on input, process and product evaluation constructs.....	159

Table 5.1	Minimum sample required	167
Table 5.2	Rate of return of questionnaires	170
Table 5.3	Skewness and kurtosis values of variables	172
Table 5.4	Range of values of skewness and kurtosis	174
Table 5.5	Profile of respondents	175
Table 5.6	Teachers' responses on input evaluation	177
Table 5.7	Teachers' responses on process evaluation (teachers' attitude)	179
Table 5.8	Teachers' responses on process evaluation	181
Table 5.9	Teachers' responses on process evaluation (challenges)	183
Table 5.10	Teachers' responses on process evaluation (school improvement)	184
Table 5.11	Teachers' responses on product evaluation	186
Table 5.12	Goodness-of-fit Indices	190
Table 5.13	Requirement for the validity of the measurement model	192
Table 5.14	Requirement for the reliability of the measurement model ...	193
Table 5.15	AMOS Output for the hypothesized model for summary statistics, variables and parameters	195
Table 5.16	Critical ratios for differences between parameters	196
Table 5.17	AMOS Output for Hypothesized model: Unstandardized Parameter Estimates	197
Table 5.18	Assessment of Normality	200
Table 5.19	The CFA results reporting for the measurement model	200
Table 5.20	Standardized Regression Weight	202
Table 5.21	Assessment of Normality	203
Table 5.22	The CFA Results reporting for the measurement model	203
Table 5.23	Standardized Regression Weights	204
Table 5.24	Assessment of Normality	205
Table 5.25	The CFA results reporting for the measurement model	205
Table 5.26	Standardized Regression Weights	207
Table 5.27	Assessment of Normality	207
Table 5.28	The CFA results reporting for the measurement model	208
Table 5.29	Standardized Regression Weights	209
Table 5.30	Assessment of Normality	209
Table 5.31	The CFA results reporting for the measurement model	210
Table 5.32	AMOS Output for the hypothesized model	211

Table 5.33	Critical ratios for differences between parameters	212
Table 5.34	AMOS Output for Hypothesized model: Modification Indices.	213
Table 5.35	AMOS Output for the Hypothesized model: Unstandardized parameter estimates	214
Table 5.36	The CFA results reporting for the measurement model	215
Table 5.37	Final characteristics of the measurement models	216
Table 5.38	Variables in the initial and final models	218
Table 5.39	Correlation between constructs	219
Table 5.40	List of Hypotheses	222
Table 5.41	AMOS Output for the hypothesized model	224
Table 5.42	AMOS Output for Model 3: Modification Indices	226
Table 5.43	AMOS Output for the hypothesized model 4	227
Table 5.44	AMOS Output for model 4: unstandardized and standardized Estimates	228
Table 5.45	Squared Multiple Correlations	228
Table 5.46	Values of fit statistics of all the four hypothesized models ...	229
Table 5.47	AMOS Output for urban model	232
Table 5.48	AMOS Output for rural model	233
Table 5.49	The moderation test using urban and rural schools	233
Table 5.50	AMOS Output for secondary model	235
Table 5.51	AMOS Output for primary model	236
Table 5.52	The moderation test using secondary and primary schools...	237
Table 5.53	Summary of Results of Hypotheses Testing	238
Table 6.1	Hypotheses and summary of results for the relationships between evaluation dimensions	253

List of figures

Figure 1.1	Dynamic Action of CIPP Evaluation Model	9
Figure 1.2	The Conceptual Framework for the SBA System Evaluation	11
Figure 2.1	National Education Assessment System	26
Figure 2.2	The Relationship between Curriculum Document and Performance Standard Document	31
Figure 3.1	The Innovation Framework of Major Areas of School Reform ...	71
Figure 3.2	Formative Assessment Model	73
Figure 3.3	A Logic Model	76
Figure 3.4	Social Constructivist assessment process: two parallel cycles..	78
Figure 3.5	Social constructivist assessment process: one dynamic system	79
Figure 5.1	Data analysis procedure	167
Figure 5.2	Hypothesized second-order model of factorial structure for the input dimension	194
Figure 5.3	Hypothesized second-order model with residual variances for second and third construct constrained equal	196
Figure 5.4	Hypothesized second-order model	198
Figure 5.5	The final re-specified model of factorial structure for input dimension of evaluation	199
Figure 5.6	The final re-specified model of factorial structure for process1 dimension	202
Figure 5.7	The final re-specified model of factorial structure for process2 dimension	204
Figure 5.8	The final re-specified model of factorial structure for process3 dimension	206
Figure 5.9	The final re-specified model of factorial structure for challenges	209
Figure 5.10	Hypothesized second-order model of factorial structure for the product dimension	211
Figure 5.11	Hypothesized second-order model of factorial structure for the product dimension	212
Figure 5.12	Re-specified model of factorial structure for product dimension.....	214
Figure 5.13	The measurement models	218
Figure 5.14	A structural model consisting 1 st -order and 2 nd -order	

constructs223

Figure 5.15 The hypothesized model - structural model 1224

Figure 5.16 The hypothesized model – structural model 4226

Figure 5.17 The final causal path with the standardized estimates229

Figure 5.18 Illustration of the moderator variable230

Figure 5.19 The hypothesized moderating model for Urban school231

Figure 5.20 The hypothesized moderating model for Rural school232

Figure 5.21 The hypothesized moderating model for Secondary school ..234

Figure 5.22 The hypothesized moderating model for Primary schools236

Definitions and Abbreviations

AaL	Assessment as Learning
AfL	Assessment For Learning
AoL	Assessment Of Learning
CA	Centre Assessment
CE	Central Examination
CFA	Confirmatory Factor Analysis
EFA	Exploratory Factor Analysis
JPN	State Education Department
MES	Malaysian Examination Syndicate
MOE	Ministry of Education
NEAS	National Education Assessment System
NPE	Natinal Philosophy of Education
PPD	District Education Department
Psi	Psychometric Assessment
PASCA	Physical Activities, Sports and Co-Curricular Assessment
PCA	Principal Component Analysis
PMR	Lower Secondary Examination
QA	Quality Assurance
SA	School Assessment
SBA	School Based Assessment
SPPBS	Management System of SBA
SPM	Malaysian Certificate of Education
STAM	Malaysian Higher Islamic Religious Certificate
STPM	Malaysian Higher School Certificate Examination
UPSR	Primary School Evaluation Test

Chapter 1: Introduction

This study evaluates the implementation of the school-based assessment (SBA) system in primary and secondary schools in Malaysia. The SBA has been introduced to Malaysia's education system since 2011. The study starts by highlighting the growing attention and interest of the shifting process of an assessment system, the conceptual framework of the study and then follows with research questions and hypotheses of the study. Chapter two discusses the education system in Malaysia and focuses more on the new assessment system called the SBA system. Chapter three provides a literature review of the assessment system, school improvement, related models and theories, program evaluation and the dimensions of evaluation. Chapter four then provides the methodology of the study. Chapter five presents the analyses and results of the study in answering the research questions proposed. Finally, Chapter six illustrates the main findings of the study, discussion of the findings, implications of the results, limitations of the study and makes suggestions for future work.



05-4506832



pustaka.upsi.edu.my



Perpustakaan TuanKu Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

1.1 Background to the Study

An education system is a fundamental aspect in building a developed nation. The structure of the current Malaysian education system was inherited from the western colonial powers and many of its characteristics reflect other Southeast Asian Ministers of Education Organization countries like Brunei, Indonesia, Philippines and Singapore (SEAMEO Secretariat, 2001). The Malaysian education system consists of pre-tertiary education (preschool, primary and secondary education) followed by tertiary or higher education. Starting at the age of five or six years old, most children go to preschool which serves as a medium in providing basic education, such as basic communication skills and English language, foster love for the country and moral values and to develop critical thinking skills amongst children (Ramlee, 2009). Primary education takes a period of six years (7 to 12 years old) with the admission age of seven years old. During the six year period, the objectives are to master the '3Ms' which are 'reading, writing and arithmetic' in the first three years, followed by the reinforcement of the '3Ms'. Pupils also need to acquire general knowledge, pre-vocational education and personality, attitude and social values

development (Ramlee, 2009). Secondary education then consists of five years of learning (13 to 17 years old), encompassing three years of lower secondary and two years of upper secondary level. Then follows a two-year period in post-secondary education either joining matriculation, technical and vocational, short term courses or Form six.

The objectives of the education system as manifested by Malaysia's National Philosophy of Education (NPE) (UMS, 2011, p.1) formulated in 1988 stated that:

Education in Malaysia is an on-going effort towards further developing the potential of individuals in a holistic and integrated manner so as to produce individuals who are intellectually, spiritually, emotionally and physically balanced and harmonious, based on a firm belief in and devotion to God. Such an effort is designed to produce Malaysian citizens who are knowledgeable and competent, who possess high moral standards, and who are responsible and capable of achieving a high level of personal well-being, as well as being able to contribute to the betterment of the family, the society and the nation at large (UMS, 2011, p.1).

This philosophy is implemented in all schools all over Malaysia. In order to achieve its objective, the Malaysian education system implements a national curriculum that aims to develop individuals in a holistic and integrated manner to produce a well-balanced community (IBE, 2011). Additionally, the five-year plan in the Tenth Malaysian Plan (2011-2015) provides a guideline in implementing the mission towards achieving a developing and high-income country as indicated by the Vision 2020 (JPM, 2010) and includes the 'Government Transformation Program [sic]' aspiration and the 'Economic New Model' to develop the country in five years ahead. In the fifth chapter of the Plan, which is to 'develop and retain world-class human capital', the stated aim is to improve the education of the people starting from early education, basic education, tertiary education until the working environment by using three core strategies:

- Reform the education system to improve students' performance
- Increase people's skills for employability, and
- Reform the labour market to make Malaysia a high income country



05-4506832



pustaka.upsi.edu.my



Perpustakaan TuanKu Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

The first core strategy, which is consistent with the interests of this study, is to reform the education system based on the NPE and to focus more on the

involvement in sports and co-curriculum. Furthermore, the education system is expected to inculcate values and ethics in achieving Vision 2020 and also to consider the use of ICT and communication to increase students' creativity, innovation and skill (JPM, 2010). Education systems around the world are also going through reforms in students' performance which is aimed at raising the bar for all students and closing the gap for lower performing groups (Fullan, 2011) and also to provide students with the competencies and higher order skills to prepare them for the challenges of the twenty-first century (Branden, 2012).

In order to improve the teaching and learning process or even to gain better impact in educational improvement, it is important to take into account the interaction between the three main aspects in education which are curriculum, instruction and assessment (Young and Giebelhaus, 2005). Higher-quality instruction could be supported by an integrated system of curriculum and assessment (Darling-Hammond and Pechione, 2010). Since assessment plays an important part, considerable attention has been placed on it by various parties. For example, lately in the United States of America, formative assessment is seen as a strategy for improvement that links the three aspects mentioned above (Clark, 2011).

1.2 Statement of the Problem

There is a large body of literature which studies the negative impact of the traditional concept of assessment on student learning. A traditional concept of assessment, which focuses most on public examination, has long been practised in the education system (Wiliam, 2001). Traditional assessment practises are not integrated into the teaching and learning process and could not give much help in improving learning or diagnosing the strength and weakness of students. The purpose of assessment has previously focused more on selection and certification for students (Wiliam, 2001) or on determining the outcome of a particular curriculum, programme or policy with little concern for any actionable information on improvement, often meant to compare with the predetermined objective (Caffrey, 2009). According to Fan (2011), students are assessed purely on their academic achievement, including knowledge and skills and are routinely practised in a time-limited situation,

tool-limited and venue-limited test. Then, most probably it will be conducted for marking and grading purposes to give report or selection process.

The traditional concept of assessment is also seen to affect students' emotion and confidence levels and to some extent could negatively influence their inner emotional strength to succeed (Stiggins, 2005). Teachers also tend to focus on those pupils they perceive as better students whom they feel have the higher chance to pass and neglect some other students, which could impact negatively on students' development (Buhagiar and Murphy, 2008). Wiliam (2001) believed that traditional assessment distort school curricula and produces results which are less reliable and valid.

Focusing on public examination as a form of assessing students has a negative rather than a positive impact on society. Public examination brings more deterioration to students and teachers as it risks producing both passive students who tend to absorb information and a passive type of teachers who tends to concentrate only on rote learning (Mercurio, 2008). It is also incapable in assessing skills like problem solving, orally expressing thoughts, school behaviour and personal and social values of students (Begum and Farooqui, 2008). Testing with multiple-choice items in most public examinations is seen to be a form of assessment that is urging students to recall and recognise discrete facts without analysing these facts critically (Darling-Hammond and McCloskey, 2008). It is also unable to evaluate pupils' broader accomplishments other than academic aspects, as with co-curriculum and sports. Public examinations therefore require teachers to focus more on examination questions rather than developing students' potential. Similarly, the Malaysian public examination is a method that orientates the public to focus on the examination as this makes up the basis for promoting students into higher level of education or for them to be awarded various scholarships (Cheah, 2010), a phenomenon which seems to deviate from the real objectives of the education system as manifested by the NPE as mentioned above. Table 1.1 lists the main features of the traditional and new concepts of classroom assessments for mathematics (Fan, 2011, p. 4).



Table 1.1: Comparison of traditional and new concept of mathematics assessment

Mathematics assessment	Traditional concept	New concept
What (content)	Cognitive domain (mainly knowledge and skill) and the results of learning	Cognitive and affective domains (knowledge, skill, ability and disposition) and both the results and process of learning
Where is it conducted	Within classrooms	Within or outside classrooms
When is it conducted	During class for a block of time	During or after class for days, weeks, months or years
How is it conducted	Conventional way (written test)	Conventional and alternative ways
Why is it conducted	Single purpose for grading and reporting learning results	Multiple purpose for improving teaching and learning
Relationship with learning	Assessment of Learning (AoL)	Assessment for Learning (AFL), Assessment of Learning (AoL) and Assessment as Learning (AaL)

Currently, the trend of the assessment system in Malaysia is changing. The Malaysian education system has started to implement SBA, an assessment system which is conducted in school and is planned, administered, scored and reported in a mannered way based on the procedures from the Malaysian Examination Syndicate (Lembaga Peperiksaan Malaysia, 2011a).

As Black (1998) believed that the terms, methods and procedures used in any assessment and testing system in a country depend on the historical and cultural background of its educational system, so it would be helpful to consider the chronology of the SBA implementation in Malaysia. The beginnings of the SBA can be traced back to the launching of the International Colloquium on the 13th to 15th of September 2005 and the Kuala Lumpur International Conference on Assessment (KLICA) on the 16th to 19th of May 2006, the latter of which used ‘Humanising Assessment’ as its theme and an idea for the SBA system in Malaysia (Ministry of Education, 2006). The

implementation of the National Education Assessment System was approved in September 2008 as a pilot project incorporating 500 schools. The follow-up meeting by the Cabinet of Ministers No. 3/2009 agreed to the Implementation of the SBA (Ministry of Education, 2011). The implementation of SBA for primary schools started in 2011 with the Year 1 students made the first move with the instruction from the MOE followed by a circular letter KP (BPSH-SPDK) 201/005/01/Jld. dated January 2011.

Since the SBA system is still in a relatively early stage of development, it seems timely to discuss some of the issues raised. Hence, a comprehensive evaluation is necessary to evaluate the system in all the dimensions previously stated (context, input, process and product). This is concordant with Gredlers' idea on programme evaluation where any educational system has to undergo careful and rigorous examination in order to improve or enhance students' educational experiences (Gredler, 1996). Mitchem *et al.* (2003) asserted that evaluation is becoming important to prevent failures in programme implementation.

1.3 The Conceptual Framework

A model or framework is a conceptual picture that shows the interrelationship between various elements involved in any given activities (Razali, 1987) and are normally presented in the form of flowcharts, web diagrams or other forms of schemata (Leshem and Trafford, 2007). Conceptual frameworks provide a theoretical clarification to support the study and a clearer picture on the purpose of the study and the process of achieving it. A conceptual framework is defined as the current version of the researcher's map of the territory being investigated (Miles and Huberman, 1984, p.33).

When developing an evaluation model to gauge the success or otherwise of a learning or assessment model, it is important to recognise there is no single evaluation design which is perfect or complete. The evaluation model does not only represent the overall evaluation framework but should also relate to the research questions or the purpose of the study (Patton, 1990). Some academics might define evaluation differently according to the purpose of evaluation. Hence, the difference in its purpose will determine the way

evaluation is conducted, the models chosen and the standards used to formulate and prosecute evaluations (Brinkerhoff *et al.*, 1983).



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

Following a comprehensive review of several evaluation frameworks and taking into consideration the definition and the purpose of evaluation in this study, the *context, input, process and product* (CIPP) Model (Stufflebeam, 1971a) is a suitable model to be used as a framework for this study, the reasons for which will be elaborated in section 3.7.6. The use of the CIPP Model is concordance to the operational definition of evaluation used in this study, which follows the definition provided by Stufflebeam and Shinkfield (1985), who stated that evaluation is a process of gathering information in guiding decision-making and quality assurance or to sum up the worth and merit of a programme. The information gained is useful to various groups of people such as the school administrations, the head teachers or the education officers. Evaluation in this study context is not aiming at collecting an interim continuous report along the implementation process and then providing stakeholders with enough information for them to revise and make improvements. Instead, it is a form of evaluation to assess a completed system to determine the system's success or failure. By using this framework, it is hoped that evaluation would help in providing guidelines to decision makers, in producing records and in the creation of concrete explanations on phenomena that happen during the programme's implementation. But, it still could turn out to be a failure if the authorities do not handle the findings in a correct manner (Stufflebeam and Shinkfield, 2007).

The conceptual framework developed in this study originated from an established model, the CIPP Model by Stufflebeam (1971a). The CIPP Model was first used back in 1965 to provide information for the purpose of decision-making and accountability through the process of evaluation of programmes, projects, products and systems. It includes four essential features which are context, input, process and product, with an ultimate aim of securing an improvement in the education field. CIPP was developed by groups of researchers that based their research from the work of the Ohio State University Evaluation Centre and the Department of Evaluation and Research in the Columbus, Ohio Public Schools.

According to Stufflebeam, evaluation involves decision-making. The context, input, process and product evaluation subsequently serve planning, structuring, implementing and recycling decisions respectively. When 'context evaluation' is involved, it gives an opportunity for the decision makers to plan the programme objectives either to confirm the present objectives, to modify the existing objectives or to develop new objectives. So, the selection of programme objectives influences planning decision. In a simple form, it is like asking oneself, "What should we do to evaluate this program [sic]?" or "Which objectives should be obtained?" (Isaac and Michael, 1982, p.10) or "Were important needs addressed?" (Stufflebeam, 2003). Next is 'input evaluation', which allows decision makers to make decisions on the structure of the programme related to strategies, personnel, resources, procedures or a prospective cost assessment in achieving the programme objectives that have been derived from planning decisions. Input evaluation involves asking questions such as, "How should we evaluate the program [sic]?" or "Which strategies or procedures should be tried?" (Isaac and Michael, 1982, p.10) or "Was the effort guided by a defensible plan and budget?" (Stufflebeam, 2003).

Making decisions on the implementation of a programme during 'process evaluation' means that decision-makers have to decide on everything related to the implementation of already selected designs, strategies or action plan; asking questions such as "Are we doing it correctly?" or "How adequately are these strategies or procedures working?" (Isaac and Michael, 1982, p.10) or "Was the service design executed competently and modified as needed?" (Stufflebeam, 2003). Lastly is the 'product evaluation'. It serves as the programme recycling decisions to determine and examine the specific outcomes of the programme, to conduct a retrospective cost assessment or cost effectiveness assessments and includes such questions as "Should the program be continued or not?" or "How effectively are the goals and objectives being accomplished?" (Isaac and Michael, 1982, p.10) or "Did the effort succeed?" (Stufflebeam, 2003). In other words, it compares outcomes of the programme with its objectives.

The relationship between the types of evaluation and decisions is shown in Figure 1.1 and Table 1.2 (Isaac and Michael, 1982). Figure 1.1 shows the dynamic action of evaluations serving the decisions in the CIPP Model. It is

‘dynamic’ in the sense that information from any stage of the decision making activity could be provided to the previous stage so that modifications on evaluations could be made. In Table 1.2, the ‘Ends’ row explains the realization of the objectives which is achieved by the planning and recycling decisions whereas the ‘Means’ row shows the act of achieving ends which is achieved by the structuring and implementing decisions (Isaac and Michael, 1982). Furthermore, decisions are also categorised under an intended act or an actual one. Planning and structuring decisions are grouped as intended ends and means respectively whereas recycling and implementing decisions are grouped as actual ends and means respectively.

Figure 1.1: Dynamic Action of the CIPP Model

	INTENTIONS	ACTUALITIES
ENDS	Planning Decisions supported by Context Evaluation <i>(What needs to be done?)</i>	Recycling Decisions supported by Product Evaluation <i>(Did it succeed?)</i>
MEANS	Structuring Decisions supported by Input Evaluation <i>(How should it be done?)</i>	Implementing Decisions supported by Process Evaluation <i>(Is it being done?)</i>