

TEACHERS' CONCEPTION OF ASSESSMENT KNOWLEDGE AND THEIR
PREFERRED ASSESSMENT SYSTEM

AMAR MA'RUF

UNIVERSITI TEKNOLOGI MALAYSIA

TEACHERS' CONCEPTION OF ASSESSMENT KNOWLEDGE AND THEIR
PREFERRED ASSESSMENT SYSTEM

AMAR MA'RUF

A thesis submitted in fulfilment of the
requirement for the award of the degree of
Doctor of Philosophy (Measurement and Evaluation)

Faculty of Education
Universiti Teknologi Malaysia

JANUARY 2018

DEDICATION

to my beloved parents, Drs. H. Muhammad Arbia Karib, St. Asiah (the late),
and Hj. Marwiyah Ramli.

to my beloved wife Syarifah Mardiana Alwy Assaggaf, S.Pd.,M.Pd.

to my children Muhammad Khidir Amar, Aisyah Shofiyyah Alkhanza Amar and
Syareefah Humairah Amar

to my the late father in law Syekh Alwy Assaggaf and my mother in law Hj.
Rosmini.

And to my Professor... Prof. Dato' Dr. Mohd. Najib Abdul Ghafar

ACKNOWLEDGEMENT

Alhamdulillah Rabbil Aalamiin, In the name of Allah, the most Gracious and the most Merciful. Praise be to Allah Subhana Wata'aala for His blessing and Rahmat that this thesis finally completed. On this particular occasion I would like to extend my deepest appreciation to number of people whose assistance contribute significantly to the completion of this thesis. I would like to present my deepest appreciation to the Governor of Sulawesi Selatan, who has given me the opportunity and scholarship to conduct this study, which I believe would provide a significant contribution to the development of education in my country, in particular Sulawesi Selatan. I wish to express my sincere appreciation to my supervisor, Prof. Dato' Dr. Mohamed Najib bin Abdul Ghafar for his guidance, invaluable time, and helpful advice that enable me to successfully complete this study. I would like to express my gratitude to Dr. Hamimah Abu Naim as the head of Measurement and Evaluation department for her support, and to Dr. Adibah Abdul Latif, Dr. A. Sukri Syamsuri, M.Hum, and Dr. Ernawati, M.Pd who validated the questionnaire of TCAQ and TPTQ instrument. I would like also to thank to the Dean of Faculty of Education, Staff of Faculty and others not possible mention here for their help. Further, I would like to say many thanks to my family, specially to my beloved wife Syarifah Mardiana Alwy Assagaf, S.Pd.,M.Pd and my beloved daughters Aisyah Shofiyyah Alkhanza Amar and Syareefah Humairah Amar for their prayer and understanding. Further, to my beloved parents Drs. H. Muhammad Arbia Karib and Hj. Marwiah Ramli, and to my mother in law Hj. Rosmini, my brother Ahsan Fatwa, ST and his family, my brothers and sisters in law who always support and pray for my success. Last but not least, to all friends in Pangsapuri Taman Desa Skudai, Dr. Andi Erham, Dr. Arfin, Dr. Hajir, Dr. Kaharuddin, Dr. Haidir and Dr. Erwin Akib, friends in G08 Graduate Space UTM for their support.

"...He hath created everything and hath meted out for it a measure".

QS. Al-Furqon (25) : 2

ABSTRACT

The shift in education system in Indonesia from centralized to decentralized system has affected the assessment system and indirectly changed the test from standardized to school-based type. Teachers' conceptions of assessment knowledge may have an influence on the changes. Theoretically, the assumption is that objectivist point of view prefers standardized test format, whilst the subjectivist view of knowledge prefers school-based test format. Thus, the study investigated the relationship between teachers' conceptions of assessment knowledge and their preferred test format. This study applied mixed method of explanatory sequential design carried out in Indonesia using multistage cluster random sampling method involving 240 teachers. The instruments used in this study were two types of questionnaires, namely "teachers' conceptions of assessment knowledge" and "teachers' preferred test format". The construct validity of the instruments was determined using the Rasch measurement model. The research findings showed that teachers' conceptions of assessment knowledge was more towards being objectivist rather than subjectivist. However, in terms of preferred test format, it was the opposite, which was inclined towards school-based rather than standardized test. This result is in conflict with the earlier assumption, that those who have conceptions of objectivist prefer the standardized test format and those with conceptions of subjectivist prefer school-based test format. Thus, in this study, teachers' conceptions of assessment knowledge did not correlate with their preferred test format. The findings show teachers' conceptions of assessment knowledge as objectivist, but they preferred school-based test. Hence, some policies need to be implemented by the government to change teachers' conceptions of assessment knowledge into conceptions of subjectivist view to be in line with the government's needs.

ABSTRAK

Peralihan sistem pendidikan di Indonesia dari sistem terpusat ke sistem desentralisasi telah mempengaruhi sistem penilaian dan secara tidak langsung mengubah dari ujian standard kepada jenis berasaskan sekolah. Konsep pengetahuan penilaian guru mungkin mempunyai pengaruh terhadap perubahan ini. Secara teorinya, andaian bahawa sudut pandangan objektif lebih mengutamakan format ujian seragam manakala pandangan subjektif berpengaruh terhadap pengetahuan menguji format ujian berasaskan sekolah. Oleh itu kajian ini mengkaji hubungan antara konsep penilaian pengetahuan guru dan format ujian pilihan mereka. Kajian ini menggunakan kaedah campuran penjelasan yang bercorak yang dilakukan di Indonesia dengan menggunakan kaedah pensampelan rawak berkelompok yang melibatkan 240 orang guru. Instrumen yang digunakan dalam kajian ini adalah dua jenis soal selidik iaitu "konsep pengetahuan guru" dan "format ujian pilihan guru". Kesahan konstruk instrumen telah ditentukan menggunakan model pengukuran Rasch. Dapatan kajian menunjukkan bahawa pengetahuan konsep penilaian guru lebih bersifat objektif dan bukan subjektif. Walau bagaimanapun, dari segi format ujian pilihan, adalah sebaliknya, iaitu cenderung kepada ujian berasaskan sekolah dan bukannya ujian standard. Dapatan ini bertentangan dengan anggapan yang lebih awal, bahawa mereka yang mempunyai konsep objektif lebih suka format ujian standard dan mereka yang mempunyai konsep subjektif akan lebih suka format ujian yang berasaskan sekolah. Oleh itu, dalam kajian ini konsep pengetahuan penilaian guru tidak dikaitkan dengan format ujian pilihan mereka. Dapatan kajian ini menunjukkan bahawa pengetahuan guru adalah objektif, tetapi mereka cenderung memilih ujian berasaskan sekolah. Oleh itu, beberapa dasar perlu dilaksanakan oleh kerajaan untuk mengubah pengetahuan konsep penilaian guru ke dalam konsep pandangan subjektif supaya selari dengan keperluan kerajaan.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	xi
	LIST OF FIGURES	xiii
	LIST OF ABBREVIATIONS	xiv
	LIST OF APPENDICES	xv
1	INTRODUCTION	1
	1.1. Introduction	1
	1.2. Background of the Problem	6
	1.3. Problem Statement	13
	1.4. Research Objective	14
	1.5. Research Question	14
	1.6. Research Hypothesis	14
	1.7. Significance of the Study	15
	1.8. Scope of the Study	16
	1.9. Theoretical Framework and Conceptual Framework	17
	1.10. Operational Definition	20

	1.10.1 Teachers' Conception of Assessment	21
	1.10.2 The Objectivist Conceptions View	22
	1.10.3 The Subjectivist Conceptions View	22
	1.10.4 The Preference of Test	23
	1.11. Summary	23
2	LITERATURE REVIEW	24
	2.1. Introduction	24
	2.2. Theories of Epistemology	24
	2.3. Epistemological Belief	26
	2.4. Types of Epistemological Belief	29
	2.5. Theories of Teachers' Conception of Assessment	30
	2.5.1. Conceptions of Objectivist	31
	2.5.2. Conceptions of Subjectivist	33
	2.6. Types of Test	36
	2.7. Standardized Test	39
	2.7.1. Negative Impact of Standardized Test.	40
	2.7.2. The Belief about Standardized Testing Test	42
	2.8. School-Based Assessment	43
	2.9. Theories of Preferences	46
	2.10. Previous Findings	47
	2.11. Summary	49
3	METHODOLOGY	50
	3.1. Introduction	50
	3.2. Research Design	50
	3.3. Population and Sampling	52
	3.4. Instrumentation	57
	3.4.1. Validity	57
	3.4.2. Reliability	64
	3.4.2.1. Summary of 100 Measured Persons	65
	3.4.2.2. Summary of 68 Measured Item	65
	3.4.2.3. Item Statistics: Misfit Order	66

3.4.2.4.	Summary of Category Structure	67
3.4.2.5.	Summary of 74 Measured Person	69
3.4.2.6.	Summary of 68 Measured Item (after deleting 26 persons)	70
3.4.2.7.	Item Statistics: Misfit Order (after deleting 26 persons)	70
3.4.2.8.	Summary of Category Structure model="r"	72
3.4.2.9.	Summary of 48 Measured Person (After Deleting 52 Persons)	73
3.4.2.10.	Summary of 68 Measured Item	74
3.4.2.11.	Item Statistics: Misfit Order	75
3.4.2.12.	Summary of Category Structure. Model="r"	76
3.5.	Data Collection	77
3.6.	Data Analysis	78
3.6.1.	Quantitative Data Analysis	78
3.6.2.	Qualitative Data Analysis	79
3.7.	Research Calendar	82
3.8.	Summary	84
4	RESEARCH FINDINGS	85
4.1	Introduction	85
4.1.1	Sample Description	85
4.2.	Objective Finding on Teacher Epistemological Belief	86
4.2.1	Teachers' Conceptions of Objectivist	86
4.2.2.	Teachers' Conceptions of Subjectivist	86
4.3.	Objective Finding on Teacher Preferred Test	88
4.3.1.	Teacher Preferred Test on Standardized Test	88
4.3.2.	Teacher Preferred Test on School-based Test	88
4.4.	The Objective Finding on The Relationship between Teachers' Epistemological Belief and Teachers' Preferred Test.	89

4.4.1.	Relationship between the conception of Objectivist and Standardized Test	90
4.4.2.	Relationship between the conception of Subjectivist and School-based Test	91
4.4.3.	Relationship between the Conception of Objectivist and School-based Test	93
4.4.4.	Relationship between the Conception of Subjectivist and Standardized Test	94
4.5.	Qualitative Analysis Finding	95
4.6.	Summary	97
5	DISCUSSION, CONCLUSION AND RECOMENDATION	98
5.1.	Introduction	98
5.2.	Discussion	98
5.2.1.	The Teachers' Conception of Assessment Knowledge in Indonesia	99
5.2.2.	The Teachers' Preferred test Types Standardized test or School based Test	101
5.2.3.	The relationship between teachers' Conception of Assessment Knowledge and teachers' preferred test system	103
5.3.	Implication	104
5.4.	Recommendation	105
5.5.	Conclusion	106
	REFERENCES	107
	Appendices A – E	118 - 176

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1.	Summaries of conceptions of Objectivist and conceptions of Subjectivist Epistemology	35
2.2.	Summaries of Standardized Test and School-Based Test	45
3.1.	Research Design	52
3.2.	The Numbers of Teachers of High School in Indonesia	53
3.3.	The Distribution of Region Republic of Indonesia	54
3.4.	Distribution of Sampling	55
3.5.	Specification of Constructing Items of Instrument of teachers' conceptions of assessment knowledge questionnaire (TCAQ).	59
3.6.	Specification of Constructing Items of Instrument of teachers' preferred test questionnaire (TPT)	62
3.7.	List of Experts panel for the validity of questionnaires	63
3.8.	Person Reliability of 100 respondents	65
3.9.	Item Reliability of 68 items of 100 persons	65
3.10.	Dimensionality table of 100 persons	66
3.11.	Fit Items table of 100 persons	66
3.12.	Scale Calibration of 100 persons	68
3.13.	Person Statistic: Misfit Order	134
3.14.	Person Reliability of 74 Persons	69
3.15.	Item Reliability of 68 items of 74 persons	70
3.16.	Item Dimensionality of 74 persons	70
3.17.	Fit Items table of 74 persons	71
3.18.	Scale Calibration of 74 persons	72

3.19.	Person Statistic: Misfit Order	138
3.20.	Person Reliability of 48 Persons	73
3.21.	Item Reliability of 68 items of 48 persons	74
3.22.	Item Dimensionality of 48 persons	74
3.23.	Fit Items table of 48 persons	75
3.24.	Scale Calibration of 48 persons	76
3.25.	Table of data analysis of the research	81
3.26.	Research Calendar of the Study	83
4.1.	The Mean Score of the Teachers' Conceptions of Objectivist	86
4.2.	The Mean Score of the teachers' Conceptions of Subjectivist	87
4.3.	Significant Different of Objectivist and Subjectivist	87
4.4	The Mean Score of the Teachers' Preferred Test of Standardized Test.	88
4.5	The Mean Score of the Teachers' Preferred Test of School Based Test	89
4.6	Significant Different between the Standardized test and School-Based Test	89
4.7	The Hypothesis testing for the Covariance between conception of OBJ and STD	91
4.8	The Hypothesis testing for the Covariance between conception of SUB and SBT	92
4.9	The Hypothesis testing for the Covariance between conception of OBJ and SBT	94
4.10	The Hypothesis testing for the Covariance between SUB and STD	95
4.11	Summary of Themes Found in Interview Analysis	96

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1.	The Research Theoretical and Conceptual Framework	19
3.1.	Explanatory Sequential Design	52
3.2.	Framework of Population	56
3.3.	Structure Calibration Scale of 100 persons	68
3.4.	Structure Calibration Scale of 74 persons	72
3.5.	Structure Calibration Scale of 48 persons	76
3.6.	Methods of data Collection	78
4.1.	The Relationship between the Conceptions of Objectivist and Standardized Test.	90
4.2.	The Relationship between the Conceptions of Subjectivist and School-Based Test.	91
4.3.	The Relationship between the Conceptions of Objectivist and School Based Test.	93
4.4.	The Relationship between the conceptions of Subjectivist and Standardized Test.	94

LIST OF ABBREVIATIONS

BAST	-	The Belief about Standardized Testing Scale
CK	-	Certain Knowledge
EB	-	Epistemological Beliefs
EBI	-	Epistemic Beliefs Inventory
EBK	-	Epistemological Belief of Knowledge
EBQ	-	Epistemological Beliefs Questionnaire
EQ	-	Epistemological Questionnaire
EBTA	-	Evaluasi Belajar Tahap Akhir
EBTANAS	-	Evaluasi Belajar Tahap Akhir Nasional
IA	-	Innate Ability
NCLB	-	No Child Left Behind
OA	-	Omniscient Authority
QL	-	Quick Learning
SAT	-	Scholastic Aptitude Test
SBA	-	School-Based Assessment
TEB	-	Teachers' Epistemological Belief
TPT	-	Teachers' Preferred Test
UAN	-	Ujian Akhir Nasional
SK	-	Simple Knowledge
UN	-	Ujian Nasional
UNESCO	-	United Nations Educational, Scientific and Cultural Organization
UUSPN	-	Undang-Undang Sistem Pendidikan Nasional
UU SISDIKNAS	-	Undang-Undang Sistem Pendidikan Nasional.

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Research Instrument	118
B	Protocol Interview	129
C	Person Statistics: Misfit Order of 100 Respondent	132
D	Person Statistics: Misfit Order of 74 Respondent	136
E	Qualitative Analysis	139

CHAPTER 1

INTRODUCTION

1.1. Introduction

The terms of measurement, assessment, and evaluation are occasionally used interchangeably. Many confusing ideas have arose due to fundamental differences among these terms as they are used in education. These terms mean very different things. Many definition have been obtained from many experts for these three fundamental terms. Kizlik (2012) defined measurement as to the process by which the attributes or dimensions of some physical object are determined. Another definition for measurement is a set of rules for assigning numbers to attributes or characteristics of people or to represent objects, traits, attributes, or behaviors (Sax, 1997; Reynolds, C.R; Livingston, R.B; Willson, V, 2010). Furthermore, measurement can be seen as a process of assigning numerals to objects, quantities, or events in other to give quantitative meaning to such qualities.

Assessment is any systematic procedure for collecting information or a process by which information is obtained that can be used to describe or better understand an issue and to make inferences about the characteristics of people or objects. Assessment is a broad term that includes testing. A test is a special form of assessment. Tests are assessments made under contrived circumstances so that they may be administered. In general, it can be said that assessment is an ongoing interactive process, in which two parties (assessor and assessee) are involved. The assessor is someone who assesses performance based on defined standards, while assessee is someone who is being

assessed. The process aims at determining the overall performance of the assessee and areas of improvement.

The term 'evaluation' is derived from the word 'value' which refers to 'usefulness of something'. When we evaluate, what we are doing is engaging in some process that is designed to provide information that will help us make a judgment about a given situation. Therefore, evaluation is an examination of something to measure its utility. Simply put, evaluation is a systematic and objective process of measuring or observing someone or something, with an aim of drawing conclusions using criteria, usually governed by set standards or by making comparisons. Evaluation is also defined as the act of collecting and providing information to enable decision-makers to function more intelligently. In terms of data analysis, evaluation includes both quantitative and qualitative data analysis and is undertaken semi-frequently. It ascertains whether the standards or goals established are met or not.

The other term that has close relationships with those three terms is test. Tests are defined as the instrument, device; a procedure of evaluation or a trial in which an individual's behavior is evaluated; and scoring using standardized procedures to know results for a certain subject that was taken by a student or a group of students (Nitko, A.J. 1996; Overton, 2012; Arikunto, 1987; Buchari, 1980; AERA, APA, & NCME, 1999). Tests also can be defined as a method to determine a student's ability to complete certain tasks or to demonstrate mastery or knowledge. Barrow and McGee (1979) defined that a test is a specific tool, procedure or technique used to obtain response from students in order to gain information that provides the basis for judgment or evaluation regarding some characteristics such as fitness, skill, knowledge, and values. Tests also can be stated as a tool, a set of questions, or an examination used to measure a particular characteristic of an individual or a group of individuals. Tests provide information regarding an individual's ability, knowledge, performance, and achievement.

Overall it can be concluded that tests are tools to acquire data, and the process of getting data is called measurement, and to value the data given is called evaluation. Sometimes the word assessment is used to give status to the data. However, there is a difference between assessment and evaluation because assessment is more on the

procedure as a whole and its system and ecosystem. Further, the practical implication of improving a system is assessment, as people assess if something is good or bad and evaluation focuses more on improving individual learning, so this thesis is about the measurements, evaluations, and tests regarding basic issues in Indonesia.

Indonesia as a developing country that has experienced rapid development. One aspect is the development of the education system in Indonesia, which has gone through two phases. The first phase was before law no. 2/1989. At this phase Indonesia only had laws for teaching and education of Law No. 4 of year 1950, Law No. 12/1954 and Law No. 22/ 1961. These laws were assumed to not reflect the education system in Indonesia, because they contradicted the constitution of 1945 and Pancasila. The second phase was Law no. 2/1989. This law was revised into the Law of National Education System no. 2/ 1989 (UUSPN, no. 2/ 1989) to create independent citizens and to develop the country. Law no. 2/ 1989 was revised in 2003 into Law no. 20/ 2003, which was recognized as UU SISDIKNAS (Undang-undang Sistem Pendidikan Nasional). The revision of Law of National Education System (UUSPN) No. 2/1989 into Education Law UU SISDIKNAS No.20/2003 was based on consideration that Law No. 2 of 1989 on the National Education System was inadequate and needed to be replaced to fit mandate changes in the Constitution of the Republic of Indonesia Year 1945 (Law of National Education System, 2003) to create a robust system for national education that can ensure the unmet needs of qualified human resources.

The education system in Indonesia has undergone rapid growth. Traditional ways of evaluation in Indonesia was through standardized tests (Ujian Nasional). The evaluation system in Indonesia was regulated by Law no. 20/ 2003 chapter XVI article 57, verse 1, which mentioned that evaluations were implemented to control education quality nationwide as a form of accountability. The focus of this verse was the standardization of evaluations in overall education levels to determine the quality of education nationally. Meanwhile, article 58 verse 1 and 2 said that:

- i. Evaluation of student learning outcomes made by educators to monitor the process, progress, and improvement of student learning outcomes on an ongoing basis.

- ii. Evaluation of learners, educational unit, and education programs conducted by the independent agency periodically, thorough, transparent, and systemic assess the achievement of national education standards.

The basic system for assessing educational output is through the testing system. The testing system is a curriculum provided by the ministry that is carried out in the school by teachers, assessed through the testing/examination system. The method for assessing outputs is called the standardized test method, which was provided by the government/ UU no. 20/ 2003 in chapter XVI article 57, verse 1 and a more school-based system as stated by UU no. 20/ 2003 in chapter XVI article 58, verse 1 and 2. The assessment is administrated through two ways. The first is through a standardized exam where all the tests are the same throughout the whole school system in Indonesia for the same level. Nevertheless, the problem for this format is that testing may be poor because the schools, experiences, and facilities are not the same. So standardized testing may not be suitable for certain schools. Second, the government came up with a new program to take care this problem, which is to school-based tests. School-based tests is where teachers are given the opportunity to set up their own items for the exam based on the needs of individual schools. Thus, there are two major approaches in assessing students, through standardized testing and through school-based testing.

The implementation of this change depends on many variables. The success of this change in implementation may be influenced by many factors such as the readiness of the school, the readiness of the teachers, the readiness of the students, the readiness of the facilities, and teaching experience. For example, are teacher's knowledgeable enough to construct their own items? Because through standardized testing they receive items from outside the school or from the panel. But in school-based testing teachers have to construct the items themselves, and if are they ready to implement school-based testing is a concern? Are the facilities and infrastructure in the school ready for the teacher to use a school-based testing? Because they still have to have be managed, and they may have to work in groups and there must be experts in the school who are knowledgeable enough in the curriculum to lead the group in constructing the items. One other variable teacher beliefs. Some teachers do what they like and follow their beliefs. Some teachers believe that knowledge is eternal and unchanging and people only have to discover. Other teachers believe that knowledge is constantly

changing and people need to catch up. Philosophically, the theory of measurement and evaluation derives from the epistemological belief that knowledge is objective and subjective.

The changing of the evaluation system in Indonesia from the standardized exam format to the school-based format influenced government policy. The success of the implementation of this policy system depends on many factors, one of which is teacher proficiency in the exam itself, such as their ability to construct items, analyze the standardized data, and use the data to explain student abilities. The other factor is teacher beliefs. There is one major factor that has not intensively investigated so far, teacher beliefs in assessment or teacher conceptions of assessment (Brown, 2004, 2006, 2008). There are many testing approaches, for example standardized tests and school-based tests. These formats derive from different conceptions of knowledge. People may look at knowledge in an objectivist way, in that is knowledge is quantitative, fixed, and unchanging. The opposite way of looking at things is the subjectivist view, which looks at knowledge qualitatively, where knowledge grows and is constructed from experience. These two approaches are derived from the epistemological of knowledge.

In relation to the assessment and evaluation system, implementing exams from standardized tests to school-based tests is derived from the conceptions of the objectivist and subjectivist views. Standardized tests are derived from the objectivist point of view, which looks at things as steady and fixed and that the majority for all people are the same. School-based exam come from the approach come from the subjectivist point view, which looks at knowledge as individualized, depending on individual perceptions, and localized because people are different.

The philosophy of measurement and evaluation is derived from the philosophy of the notion, which is called epistemology. Epistemology is divided into two approaches, objectivist and subjectivist. Both objectivist and subjectivist approaches are called conceptions in regard to assessment. The objectivist refers to quantitative data where the data is the same and constant for everyone, making it standardized. The other conception of assessment is the subjectivist view. The subjectivist view refers to the qualitative part of knowledge which is individualized and constructed by different

individuals. For example, in Indonesia what should be learnt in Sumatra may be different from Kalimantan because of their different contexts and areas.

However, the changing of standardized tests to school-based tests by the government cannot be separated from the success of decentralized systems that shifted all decisions from the central government to local or province governments, and this may influence teacher conceptions. Teacher conceptions that the true nature of knowledge is objectivist may cause them to prefer standardized tests. On the other hand, if a teacher view on knowledge is subjectivist they may prefer school-based tests. Thus, there is an assumption that centralized to decentralized system teachers' should be able to implement school-based tests, but in Indonesia there have been some problems in the evaluation system. There was a report stating that besides the lack of facilities and infrastructure in some rural areas in Indonesia, teachers were not ready to implement school-based tests as they have little experience in constructing their own tests (Syahril, 2007; Ali, 2014). This is what this thesis would like to investigate.

1.2. Background of the Problem

Education is a form of learning in which knowledge, skills, and habits were transferred from one generation to the next through teaching, training, or research. Education is often under the guidance of others, but may also be autodidact (Dewey, 1942). Education is described as normative or prescriptive in philosophy or descriptive in science. From the philosophical point of view education is a field of applied philosophy that examined aims, forms, methods, and results as both a process and a field of study (Kneller, 1964; Dolhenty, 2010; Frankena, et al, 2002; Philips, 2008). Education in terms of normative philosophy deals with goals, norms, and standard for conducting the process of education. Education in terms of descriptive science provides a hypothesis or a set of hypothesizes that have to be verified by observation and experiment (Kneller, 1964). In other words, normative educational theory or philosophy provides goals for education, whereas descriptive theory or science provides concrete data that would help realize the goals suggested by the philosopher.

Philosophy and education are two different things that relate to each other in terms of human nature. It is a comprehensive system of ideas about human nature and the nature of reality. Furthermore, philosophy refers to the basic concepts, beliefs, and attitudes of an individual or group. It is also distinguished from other ways of addressing such problems by its critical, generally systematic approach and reliance on rational argument (Ayn Rand, 1982; Teichmann, Jenny and Evans, Katherine C, 1999; Grayling, A.C, 1995; Quinton, 1995). Since philosophy addresses a basic and pervasive issue, it can be used as a guide for life to determine the course that people should take in life and how people should treat each other. Therefore, it can be said that almost all human life is affected and governed by philosophical considerations. Many experts considered philosophy as the mother of all the sciences and the oldest discipline. Philosophy is the root of all knowledge.

The implementation of education should be properly managed. Well-managed education is today generally paid for and almost entirely governed by public bodies or private institutions. Government as the policy maker is responsible to make laws and rules that regulate the operation of the education system (Friedman, 1962; Bell & Stephenson, 2006). Education cannot be separated from the government in terms of formal education. As other countries in the world, education development in Indonesia has experienced several phases including phase 1 before Law No. 2/1989, phase 2 after Law No. 2/ 1989, and phase 3 Law No. 20/ 2003 (SISDIKNAS). Before 1999, Indonesian education is centralized and all decisions were made by the central government. The managerial system consisted of an extended hierarchical structure of national, regional or provincial, district and sub-district levels. This centralized system included the central government decided policies of teacher recruitment, retention, promotion, and management. The provincial level implemented these policies without consideration for local conditions. District and sub district levels implemented both national and provincial policies into school practices. In this situation the principals at the foot of the hierarchy had almost no power to decide teacher allotment and promotion. (UNESCO, 2005).

However, after Law No. 22. 1999, the centralized system was abolished along with the hierarchical relationships between districts/municipals, provincial, and central administration. Law No, 22. 1999 stated that Local Governance was the legal basis for

the decentralization of authority in education. This law was articulated in Government Regulation No. 25 of 2000, which sets the shared authority between central and district governments. According to this regulation, the more operational and technical arrangements in educational implementation belong to the district or local government. The only authority left to the central government were those related to the setting of national policies for competency standards, national curriculums, education calendars, and evaluation. (UNESCO, 2005). Hence, authority previously located at the central government shifted largely into the hands of local governments.

The decentralized system was applied after the law no 22/1999. The central government and local governments had a part in the education system, where the central government handled the setting of national policies for competency standards, national curriculums, education calendars, and evaluation. The local government focused on operational and technical educational implementation. The implementation of a decentralized system effected educational aspects such as instruction, curriculum and syllabus, educational calendar, and evaluation. Thus, the central government referred to policy at the level of legislation and regulations, while the local government handled operational and technical implementation.

One of the aspects that influenced the decentralization system in Indonesia was assessment. Assessment development in Indonesia experienced changes in format several times. The following were the changes in the assessment format.

- i. 1950 – 1960s: in this period exams are called Final Exams (Ujian Penghabisan). Final exams were held nationwide and all testing questions were created by the Ministry of Education and Culture. All questions are in an essay format. The results of these tests were not checked at the school but in the district center.
- ii. 1965 – 1971: in this period all subjects were tested in one program called State Exams (Ujian Negara). Exam materials were created by the central government and were valid for all of Indonesia. Exam times were also determined by the central government.

- iii. 1972 – 1979: the government allowed the schools or group of schools to hold their own exams. The test and assessment processes were conducted in each school or group. The government just released guidelines.
- iv. 1980 – 2001: the national final exam called the National Final Learning Evaluation (EBTANAS) was organized. This final test used EBTANAS for several main subjects. Other subjects that include EBTANAS or non-Ebtanas were tested in the (EBTA) Final Test Evaluation. EBTANAS was conducted by the central government (Standardized test). EBTA was conducted by the local or district government (School Based Test). Graduation was determined by a combination of two EBTANAS and EBTA evaluation exams plus daily values listed on the report card. In EBTANAS the students passed if the average value of all subjects was six.
- v. 2002 – 2004: EBTANAS was replaced by national assessments of learning outcomes and changed into a National Final Examination (UAN). UAN graduation in 2002 was determined by the value of individual subjects. The test for national final examination was prepared by the Ministry of Education (Standardized Test) and the school cannot change the score for UAN.
- vi. 2005 – Present: The government still conducts similar tests, although there was a changing in name from UAN (National Final Examination) to UN (National examination). Despite heavy criticisms for UAN, UN still uses the same format.

Based on the above periods, it can be observed that during 1950 - 1975 Indonesia implemented standardized testing for almost all subjects for all students at the end of elementary, middle school, and high school. School based tests or non-standardized tests have been implemented since 1975, where schools are given the authority to design and manage the final exam based on guidelines from the central government (Furlong, 2004). In 1980 Indonesia again used standardized tests or centralized exam tests, which was called EBTANAS, although during this period of time schools still had the authority to conduct EBTA (Final Learning Evaluation). In EBTANAS, decisions about student graduation were largely in the hands of the schools. Thus, students who performed poorly in these EBTANAS tests were still able to graduate provided they performed well in school. EBTANAS scores were just part of a total scoring component for student graduation, besides EBTA (Final Learning

Evaluation), and school grades. Furthermore, EBTANAS was changed in 2002 to become UAN, where the test was made by the Ministry of Education in standardized test form. This centralized system is currently used today.

The implementation of standardized tests has many advantages. Standardized tests were objective, valid, and reliable. Standardized tests were objective since they were not open to bias and their interpretation can be simplified by statistical figures. They were valid because the information collected was trustworthy and reliable since the results of the test were consistent. (Oakes and Lipton, 2007). Cizek (2001) highlighted five positive consequences of Standardized tests. A improvement in the quality and focus of the professional development of educators, an awareness of the needs of all students, an increase in the numbers of assessment literate teachers, an increase in data-driven decision-making, and an increase in the quality of tests. Furthermore, standardized testing has assumed a prominent role in recent efforts to improve the quality of education (Herman, J.L and Gholan S, 1991).

However, among the advantages and the benefits of standardized tests, there are many studies which found that standardized tests have disadvantages and a negative impact on education. The validity and value of traditional standardized tests are subjects of increasing debate. (Cannell et al., 1987) raised the question of whether improvements in test score performance actually signaled improvements in learning. Other studies found that standardized tests tend to narrow teaching content, create mismatches between curriculum and classroom instruction, neglect higher thinking skills, and use a meaningless and irrelevant multiple choice format. (Baker, 1989; Herman, 1989; Shepard, 1990). In line with the above opinions, standardized tests emphasize the need to focus only on content that is tested. Hence, the curriculum becomes increasingly narrowed in classrooms, including that which emphasizes higher-order thinking skills (Amrein & Berliner, 2002; Birkmire, 1993; Darling-Hammond, 1991; Madaus, 1988; Pedulla, 2003). Oakes and Lipton (2007) criticisms on standardized testing has been commonly ignored by policymakers. They argued that there were three weaknesses in standardized testing. First, learning theories in behavioral psychology from the nineteenth-century did not reflect cognitive and educational viewpoints. Second, it inherited flaws in the logic and technology of IQ test where tests were designed to produce a wide range of scores, with most of the test-

takers score in the average range and only few reaching high scores. Third, it is culturally biased, since different areas have different cultural strengths that were emphasized.

Due to the disadvantages of standardized tests, it is necessary to identify teachers' conceptions of assessment knowledge to discover their preference. Previous research has shown that teachers generally report dissatisfaction with testing, skepticism on their validity or usefulness, and a feeling that tests rob them of their ability to shape the curriculum. (Green & Stager, 1986; Haladyna et al., 1991; Nolen et al., 1992). On the other hand, teachers teach based on the test when they know that important decisions would be made based on test scores. This has led some teachers to ignore teaching material based on the curriculum, redefine course objectives, and resequencing course content in an attempt to improve test scores, which discourages teachers from using team teaching approaches and from changing their methods to facilitate serious student learning. (Madaus., 1988; Smith., 1989; Corbett and Wilson., 1988; Stodolsky., 1988).

Associated with some disadvantages of standardized testing above, there is a need to find an alternative test model in order to solve testing problems. Gipps (1999) explained that the focus of assessment has shifted to a broader assessment of learning, individual enhancement, engagement with students, and teacher involvement in the assessment process. Due to this change school based assessment was introduced. Izard (2001); Raivoce and Pongi (2001); and Grima (2003) perceived School-based assessment as the process of collecting evidence on what students have achieved in important learning outcomes that do not easily lend themselves to pen and paper tests. A school-based test model is one alternative test model that is able to solve this problem. According to Raffan (2001) school based assessment provides teachers with documents such as advice booklets, videos, training meetings, and assessor networks. Furthermore, this assessment also gave sufficient time and attention to moderation procedures. Portal (2003) gave advice on the importance of School based assessment, as it was not important to use a standardized test or national test in order to be able to deliver a clear picture of individual achievement, unless an assessment practice effects public confidence.

Both standardized testing and school based assessment are different epistemologically. Standardized tests are objective. Magee, (2012) stated that objectivity is related to the conceptions of the objectivist model. The objectivist model is concerned with accountability, efficiency, and quality control. Its evaluation information was considered to be “scientifically objective”, where objectivity was achieved by using objective instruments like tests or questionnaires. Presumably, results produced with these instruments were reproducible. Data was analyzed by quantitative techniques that were also “objective”. The analyzed data can be verified by logical inspection regardless who uses the techniques.

School-based assessment is an assessment format where teachers are given greater responsibility in developing assessments and linking them to effective learning. Student achievements would be judged and graded by teachers based on the criteria and standards specified in subject syllabuses, and were moderated by review panels consisting of subject matter experts (Queensland Studies Authority, 2007). Since teachers are the most responsible part of this assessment, subjective influence might occur. Subjectivist models were concerned with empirical experience and tacit or implicit knowledge. Validity depends on relevance of evaluator backgrounds. Procedures were based on evaluator perception or experience and cannot be reassessed by other people. Basically, subjectivism emphasized community approval. (Mohamed Najib, 2009). Another meaning of subjectivist is something based on personal judgment and personal desires. Each person is the best judge of events for themselves. The subjectivist evaluator is more concerned in their work with specific casual statements.

Therefore, the writer saw that changing from a centralized to decentralized system in Indonesia affected many aspects, including the assessment system, such as the implementation of standardized tests to school-based tests. This change also creates some problems in the implementation of school-based tests. As mentioned earlier teachers’ either have an objectivist view or an subjectivist view. These two views may influence the assessment format that they prefer. Since there was very little research on this topic, this led the writer to investigate whether or not there was a relationship between teacher conceptions in terms assessment knowledge towards preferred assessment formats.

1.3. Problem Statement

According to the UNESCO report (2005), in the transition from New Order rule in 1999 to the present, the national government of Indonesia launched radical democratic policies in electoral politics, governance, and education. The decentralization policies include transferring responsibility for core functions to Indonesia's district governments. One aspect of decentralization is education, which is mandated by the national five year program for school-based management (Law No. 22 of 1999). This law was articulated in Government Regulation No. 25 of 2000, which established shared authority between the central and district governments. As a result of this decentralization policy, there was a trend from standardized tests to school-based tests. Thus, teachers should equip themselves with the necessary skills and attitudes so that the government's ideals will be realized in the future.

Teachers play an important role in teaching, including the evaluation process. Therefore, as argued earlier teacher conceptions on assessment knowledge can be divided into objectivist and subjectivist conceptions. Teacher conceptions on assessment can be influenced by the type of test, because objectivists contain many elements that can be related to standardized tests and subjectivists contain many elements that can be related to school-based tests.

It is important to identify teacher conceptions on assessment knowledge in relation to their preferred test format. It is assumed that teachers with fixed views on the world tend to be objectivist and prefer a standardized test format. Meanwhile, teachers who have an dynamic view of the world tend to be subjectivist and prefer a school-based test format. The Indonesia educational system has changed from centralized to decentralized. Thus, the main test format is the school-based test format. If teachers still like this standardized test format, there will problems in the implementation of a decentralized system. Hence, the main focus of this study is to investigate teacher conceptions on assessment knowledge in relationship with their preferred test system, which would facilitate the transformation of standardized tests into school-based tests.

1.4. Research Objective

This studies research objectives are the following:

- i. Identify teachers' conception of assessment knowledge among Indonesian high school teachers.
- ii. Identify teachers' preferred assessment formats among Indonesian high school teachers.
- iii. Investigate the relationship between teacher' conception of assessment knowledge and teacher preferred assessment formats.

1.5. Research Question

The research questions are:

- i. What were teachers' conception of assessment knowledge among Indonesian high school teachers?
- ii. What were the teachers' preferred assessment formats among Indonesian high school teachers?
- iii. Was there any relationship between teacher conceptions of assessment knowledge and teacher preferred assessment formats?
- iv. How was the teachers' conception of assessment knowledge influence their preferred assessment format?

1.6. Research Hypothesis

Based on the objectives, there was one main hypothesis and four sub hypotheses as follows:

- Ho₁ : There was no significant difference between teachers' conception of assessment knowledge and teacher preferred test formats.
- Ho_{1.1} : There was no significant relationship between objectivists and standardized tests.
- Ho_{1.2} : There was no significant relationship between subjectivists and school-based tests.
- Ho_{1.3} : There was no significant relationship between objectivists and school-based tests.
- Ho_{1.4} : There was no significant relationship between subjectivists and standardized test.

1.7. Significance of the Study

This study is expected to promote the development of education in Indonesia. For students as the target of test implementation, this study is expected to develop good assessments that can measure student abilities in terms of learning targets, and standardized nation education levels. In terms of teachers, this study is expected to provide new perspective on teacher conceptions of assessment knowledge in relation to their preferred test format.

This study is expected to contribute to schools since teachers that have good perspective on assessment knowledge would influence teacher performance in the learning process. For example, teachers would have a good preparation before, during, and after the learning process in the classroom. This study also provides information on the relationship between teacher conceptions on assessment knowledge and preferred test formats, which provides a new perspective on teaching and learning.

For the government as the policy maker, the results of this study are expected to give alternative views on test formats conducted in Indonesia. The government also is provided activities that help improve teacher capabilities for designing and constructing their own tests in relation with the change from standardized tests to school-based tests. In terms of decentralization, the central government is required to

provide opportunities for local government to make their own decisions on the assessment and evaluation process.

1.8. Scope of the Study

The variable of this research focused only to teacher conceptions of assessment knowledge and their preferred tests. The researcher only conducted mixed methods in explanatory sequential design. The population of this research is teachers in senior high schools in Indonesia. The researcher conducted a multistage cluster random sampling method since population and sample of this study were spread out.

In line with objectives of this study, the researcher developed two instruments. The first instrument is a close-ended questionnaire called the Teachers' Conceptions of Assessment Questionnaire (TCAQ) and the Teachers' Preferred Test Questionnaire (TPTQ). They were developed during the process of determining the study objectives, which were obtained from a literal review of the research variables. The second instrument is an open-ended protocol interview to provide supporting evidence for the two variables.

The validity of the instrument was tested by using a literature review by defining concept, construct, and operational definitions (Mohd Najib Ghafar, 2011; Fraenkel, Wallen, and Hyun, 2012; Johnson & Christensen, 2012). A panel of experts was also requested to identify the validity of the instrument by providing suggestions and recommendations to improve the sentence structure of each item of the instrument. Expert reviews of the content was one of the most important tests for content validity. (Cohen and Swerdilk, 2002; Mohd Najib Ghafar, 2003). In order to prove the validity and reliability of the instrument, the researcher administered a pilot study. The reliability of the instrument was analyzed using the RASCH Model to find out whether or not some items needed to be deleted or modified. The open-ended protocol interview was developed by the researcher following the study objectives to support the close-ended questionnaire.

The collection of data in this research used an inter-method mixing, which according to Johnson and Christensen (2012) is when two or more methods of data collection are used in a research study. The researcher acquired data from two instruments the close-ended Teachers' Conceptions Of Assessment Questionnaire (TCAQ) and Teachers' Preferred Test (TPT) questionnaire, and an open-ended protocol interview. Data analysis in this study used mixed data analysis, where both quantitative and qualitative analyses were used.

1.9. Theoretical Framework and Conceptual Framework

A theoretical framework serves as the lens by which a researcher observes a particular aspect to examine a topic of a subject field. In other words, it clarifies or explains the rationale, justification, or basis of a study (Khan, 2010). A theoretical framework is a theoretical perspective. It can be just a theory or more general approach to understand something. A theoretical framework is used to design a theoretical structure prior to research construction and it introduces and describes the theory that explains the problem being researched (Leedy, 1974; Torraco, 1997; William, 2008).

A theoretical framework must consist of the following elements (Gregory Herek, 1995):

- i. An explicit statement of the hypothesis or theoretical assumptions on which the research was based and the relevant research method that guided the researcher in his or her attempt to test the assumption - the why and how of the research.
- ii. To what extent did the research build upon existing research or knowledge, or a clear explanation of how the hypotheses connected the researcher to existing knowledge (the literature review).
- iii. A clear articulation of the theoretical assumption on which the research was established, and how it permitted the researcher to move from simply

explaining a phenomenon to generalizing about various aspects of that phenomenon through observation.

- iv. A comprehensive description of the research method to be used and how it continued from a theoretical hypothesis or theory to an empirical hypothesis or theory.

In this research, the theoretical and conceptual framework involved the variable the research aimed to investigate. The theory is based on constructivism (Piaget, Vygotsky and Bruner) as the general theory. Based on this general theory, the researcher generated two basic concepts, the theory of epistemological belief and the concept of preference. The epistemological belief was then divided into objectivist and subjectivist, which consisted of ten constructs as the independent variable. The theory of preference consisted of five constructs for both standardized tests and school-based tests as the dependent variables. Figure 1.1 shows the theoretical and conceptual frameworks.

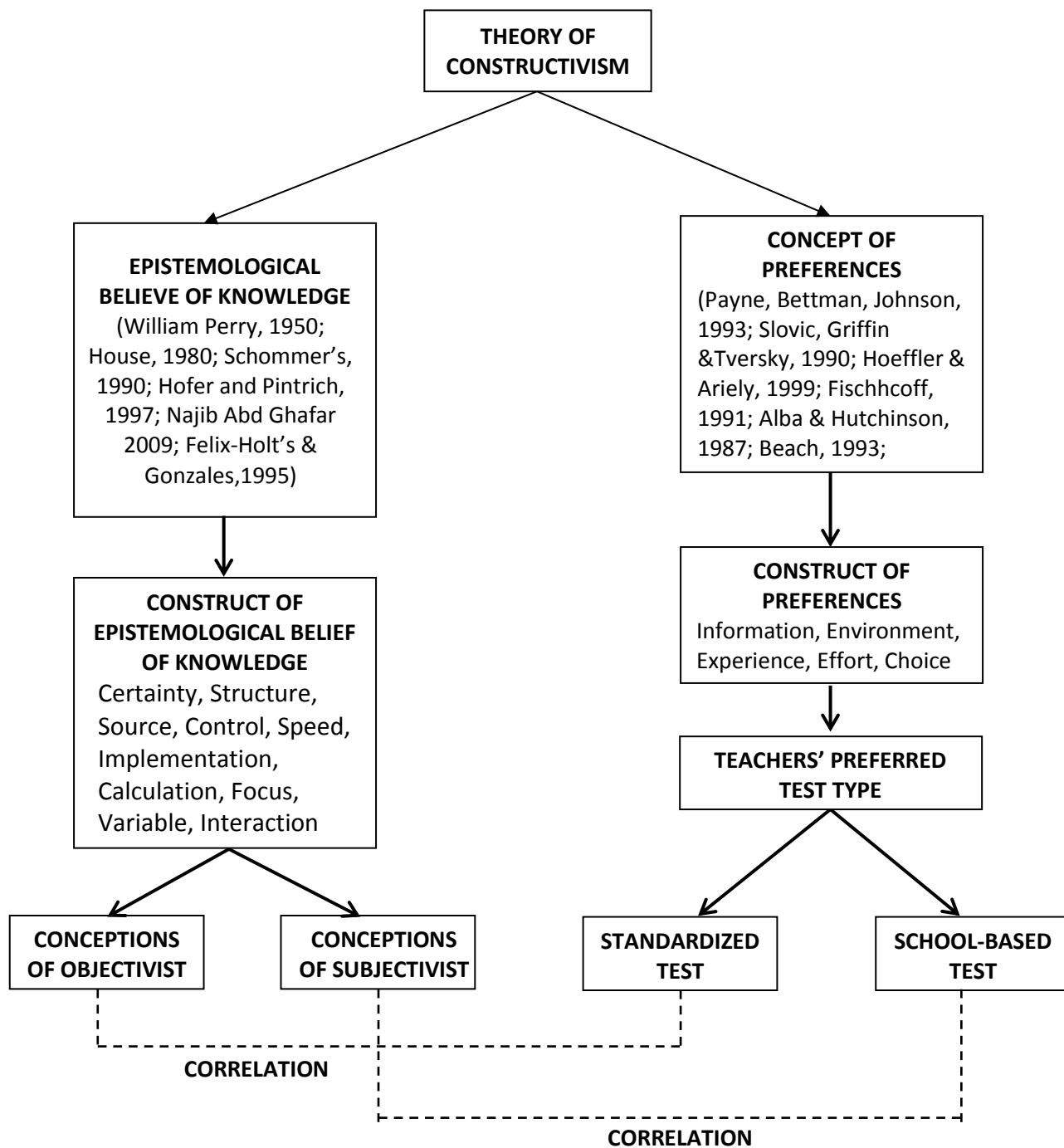


Figure 1.1: The Research Theoretical and Conceptual Framework

1.10. Operational Definition

William Perry (1970) studied epistemological belief. He focused on developmental stages and suggested that “personal epistemology was unidimensional and develops in a fixed progression of stages”. According to Schommer (1990) however, epistemological beliefs were too complex to explain in a single dimension and she defined personal epistemology as “a belief system that was composed of several more or less independent dimensions”. She introduced a system of belief about knowledge acquisition called Epistemological Beliefs (EB) that is multidimensional. She proposed five dimensions, which consisted of Quick Learning (QL), which relates whether the learning process is a fast, all-at-once process or a slow and gradual process. The next dimension was Certain Knowledge (CK), which measured whether or not knowledge, once acquired, is firmly established or is something that changes. Simple Knowledge (SK) captured whether knowledge is spread or interrelated and integrated at the level of abstract concepts. Innate Ability (IA) indicates whether or not the learning ability is genetically predetermined or acquired by experience. The last dimension was Omniscient Authority (OA), which covered whether or not the knowledge is transferred through experts or obtained by personal reasoning and the observation of empirical evidence.

Hofer and Pintrich (1997) classified the dimensions of personal epistemology into two areas: the nature of knowledge and the nature or process of knowing. They classified certainty of knowledge and simplicity of knowledge as the nature of knowledge and source of knowledge and justification of knowledge as the process of knowing. Subsequently they focused on the factors that affected students’ epistemological beliefs, such as age, gender, socioeconomic status, parents’ educational level, achievement, learning approaches, learning environment, and motivation. (Hofer and Pintrich in Fatma Kurt, 2009).

1.10.1. Teachers' Conception of Assessment

The study of teachers' conception of assessment is a critical issue in the field of assessment research. However, the domain of teacher belief is a very attractive one because there was clearly evidence that beliefs about teaching, learning and curricula strongly influence how teachers teach and what students learn or achieve. (Clark & Peterson, 1986; Pajares, 1992; Savasci-Acikalın, 2009). According to this new perspective, teachers' conception or beliefs are key factors regarded as essential determinants of instructional activity and of student learning processes. This is because teachers have a variety of beliefs that influence the quality of their performance. Hence, they develop their own beliefs about epistemology, their students, the content of teaching materials, how they teach, and social aspects related to teaching. (Levin, 2015). This belief functions as a filter that affects their method of implementation. According to Jeppe Skot (2015), teacher beliefs are individual mental constructs that are value laden and are subjectively true. They are relatively stable toward significant social experiences and have an increased impact over teacher interpretations and contributions in the context of their teaching.

In the research of teacher beliefs, if we refer to teacher's beliefs on assessment, the term frequently used is "conceptions". The term conception was initially introduced by Thompson (1992), and it refers to a general mental structure, encompassing beliefs, meanings, concepts, propositions, rules, mental images, preferences, and the like. National assessment conceptions describe overall teacher perceptions and their assessment awareness (Barnes, Fives & Dacey, 2015). Brown (2004, 2006, and 2008) used the term teacher conceptions in relation to assessment. In the specific research area of teacher beliefs in relation to assessment the established term is that of conception.

In relation with assessment, teacher conceptions are categorized into four categories (Brown, 2004, 2006; Davis & Neizel, 2011; Haris & Brown, 2009; Remesal, 2007). These categories are briefly described through the following assumptions: a. assessment improves teaching and learning; b. assessment holds the students responsible for their own learning; c. assessment charges institutions and

teachers with the responsibility of teaching students/pupils; d. assessment is irrelevant if it negatively affects teachers, students/pupils, curricula, and teaching.

1.10.2. The Objectivist Conceptions View

Objectivist conceptions according to this study are absolute, static, state knowledge as facts, are acquired from experts through objective data, believe that an individual's ability to learn is fixed at birth, knowledge is a concrete knowable fact, knowledge needs a predetermined amount of time to be acquired, is performed by an administrator, uses a quantity based scoring intensity, requires society agreement, data can be manipulated, and that data can be generalized. (Schommer, 1990, 1994; Hofer & Pintrich, 1997; Mohd Najib Ghafar, 2009; House, 1980; Felix-Holt's, 1995; Felix-Holt & Gonzalez, 1999).

1.10.3. The Subjectivist Conceptions View

The Subjectivist conception is seen as relative, tentative, changing over time, naturalistic, artistic, non-traditional, holistic, descriptive, ethnographic, is represented by complex theories, is acquired from social constructions through subjective measurements, believes that the ability to learn changes throughout an individual's life, knowledge is explained as relative, contingent, contextual, believes that most things can be learned by most people if enough time is dedicated to it, learning is a gradual process, is performed by judges or panels, is personally descriptive, is quality-based, requires individual agreement, that data cannot be manipulated, and that data cannot be generalized. (Schommer, 1990, 1994; Hofer & Pintrich, 1997; Mohd Najib Ghafar, 2009; House, 1980; Felix-Holt's, 1995; Felix-Holt & Gonzalez, 1999).

1.10.4. The Preference of Test

This study defined the preference between standardized tests and school-based tests as based on information available at preference elicitation. Individual preferences often depend on the decision-making environment. Individual experiences are the foundation of their preference structures, and the processes associated with such experiences lead to preferences that stabilize over time. Effort is simply the amount of mental energy that consumers/individuals invest in making up their minds. Choice was conceptualized as a process by which preferences were consolidated to arrive at a resolution for a chosen task. (Payne, Bettman, Johnson, 1993; Slovic, Griffin & Tversky, 1990; Hoeffler & Ariely 1999; Fischhoff, 1991; Alba & Hutchinson, 1987; Beach, 1993; Montgomery, 1983).

1.11. Summary

This chapter covered the main issue facing the education system in Indonesia. One important aspect was the assessment system. Assessment systems in Indonesia have been developed since the independence of Indonesia to the present day. It has shift from centralized to decentralized over time. Teachers are an component involved in the assessment process that was assumed to know their epistemological beliefs toward their preferred test type. In line with that, the objective of this study described was to develop close-ended Teachers' Conceptions Of Assessment Questionnaire (TCAQ) and Teachers' Preferred Test Questionnaire (TPTQ) and an open-ended protocol interview.

REFERENCES

- Alba, J. W., & Hutchinson, W. (1987). *Dimensions of Consumers Expertise*. Journal of Consumer Research, 13, 411 – 442.
- Alhojailan, M. (2012). *Identification of learners' attitudes regarding the implementation of read/write web, blog tools: a case study in higher education*. 7th Disco conference reader: New media and education. In: Prague: Centre for Higher Education Studies, pp. 58–73
- Amrein, A. L., & Berliner, D. C. (2002). *An Analysis of Some Unintended and Negative Consequences of high-stakes testing*. Tempe, AZ: Educational Policy Studies Laboratory, Arizona State University. Retrieved June 28, 2004 from [http:// www.asu.edu/educ/eps/EPRU/documents/EPS:L-0211-125-EPRU.pdf](http://www.asu.edu/educ/eps/EPRU/documents/EPS:L-0211-125-EPRU.pdf).
- Anastasia, A., & Urbina, S. (1997). *Psychological Testing (7th Ed)*. Upper Saddle River, NJ: Prentice Hall.
- Anderson, Scarvia B. (1975). *Encyclopedia of Educational Objectives*. Jossey-Bass Publisher San Francisco. California, USA.
- Arikunto, Suharsimi. (1987). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta, Indonesia. Bina Aksara.
- Aypay, Ayşe. (2011). *The Adaptation of the Teaching-Learning Conceptions Questionnaire and Its Relationships with Epistemological Beliefs*. Eğitim Danışmanlığı ve Araştırmaları İletişim Hizmetleri Tic.Ltd. Şti.
- Baker, E.L. (1989). *Can We Fairly Measure the Quality of Education?* (CSE Tech. Rep. No.290). Los Angeles, CA: University of California, Center for Research on Evaluation, Standards, and Student Testing (Grant No.OERI-G-86-0003).
- Barrow H.M. and McGee R. (1979). *A Practical Approach to Measurement in Physical Education*. Lea & Febiger, Philadelphia. U.S.A.

- Bauer, J., Festner, D., Gruber, H., Harteis, C., & Heid, H. (2004). The effects of epistemological beliefs on work-place learning. *Journal of Workplace Learning*, 5 (6), 284–292.
- Baxter Magolda, M.B. (1992). *Knowing and Reasoning in College: Gender-Related Patterns in Students' Intellectual Development*. San Francisco: Jossey-Bass Publishers.
- Beach, L. R. (1993). *Broadening the Definition of Decision Making: The Role of Prechoice Screening of Options*. *Psychological Science*, 4, 215-220.
- Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986). *Women's Ways of Knowing: The Development of Self, Voice, and Mind*. New York: Basic Books, Inc., Publishers
- Bell, Les & Stephenson, Howard. (2006). *Education Policy: Process, Theme, and Impacts*. London.Routledge.
- Birkmire, K.J. (1993). *Ninth Grade Proficiency, Is it the best Measure of Learning: A Survey of Teacher Attitudes*. ERIC Document Reproduction Services No. ED 361 406
- Bond, T.G. & Fox, C.M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences*. 2nd Edn. New Jersey. Lawrence Erlbaum Associate Inc. Publisher.
- Bracey, G.W. (1989). *The \$150 Million Redundancy*. *Phi Delta Kappan* , 70 (9), 698-702.
- Broadfoot, P. (1995) *Performance Assessment Perspective: International trends and Current English Experience*. In H. Torrance (Ed) *Evaluating Authentic Assessment*. London: The Palmer Press.
- Brown, Gavin. (2011). *School Based Assessment Methods – Development and Implementation*. Paper. The ACER Australian Council for Education. Australia.
- Bruning, Schraw, Norby, and Ronning (2004). *Cognitive Psychology and Instruction*. Paper. London. Pearson; 4th edition.
- Bryman, A and Cramer, D. (1990). *Quantitative Data Analysis for Social Scientists*. London. Routledge.
- Buchari, M. (1980). *Tehnik-Tehnik Dalam Evaluasi Pendidikan*. Bandung. Indonesia.
- Burns, Robert B. (1994). *Introduction to Research Methods 2nd Ed*. Melbourne. Longman Cheshire.

- Burns, N. & Grove, S. K. (2003). *Understanding Nursing Research*. (3rd Ed). Philadelphia, PA: W. B. Saunders
- Cannell, J.J. (1987). *Nationally Normed Elementary Achievement Testing in America's Public Schools: How all 50 states are above the national average*. *Educational Measurement: Issues and Practice* , 7(4), 12-15.
- Cizek, G. J. (2001). *More Unintended Consequences of High-Stakes Testing*. *Educational Measurement*, 20(4), 19 – 27.
- Clarke, A. (1999). *Evaluation Research: An Introduction to Principles, Methods and Practice*. London. Sage Publication.
- Creswell, J. W. & Plano Clark, V. L. (2007). *Designing and Conducting Mixed Method Research*. Thousands Oaks, CA: Sage.
- Creswell, J.W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. (3rd Edition). United States of America. Sage Publication, Inc.
- Cohen, L, Manion, L., & Morrison, K. (2000). *Research Methods in Education* (5th Ed). London. Routledge.
- Cohen, Ronald. J & Swerdilk, Mark. E. (2002). *Psychological Testing and Assessment: An Introduction to Test and Measurement*. New York. McGraw-Hill Higher Education.
- Cronbach, L.J. (1990). *Essentials of Psychological Testing* (5th Ed). New York. Harper Collins.
- Crotty, M. (1998). *The Foundations of the Social Research. Meaning and Perspective in the Research Process*. London. Sage Publications.
- Darling-Hammond, L. (1991). *The Implications of Testing Policy for Quality and Equality*. *Phi Delta Kappan*, 73(3), 220 – 225.
- Dewey, John (1942). *Democracy and Education: An Introduction to the Philosophy of Education*. North Caroline. Hayes Barton Press.
- D. C. Phillips. (2008). *What is Philosophy of Education*, in Sage Handbook of Philosophy of Education.
- Dolhenty, Jonathan. (2010). *Philosophy of Education and Wittgenstein's Concept of Language-Games*. The Radical Academy. Retrieved 19 November
- Dorr-Bremme, D., & Herman, J. (1986). *Assessing Student Achievement: A Profile of Classroom Practices*. CSE Monograph Series in Evaluation, 11. Los Angeles, CA: Center for the Study of Evaluation.

- Ebel, Robert. (1972). *Essential Educational Measurements*. Prentice Hall, Inc. Englewood, Cliff. New York, USA.
- Ernest, P. (1995). *The One and the Many*. In L. P. Steffe & J. E. Gale (Eds.), *Constructivism in Education*. NJ: Lawrence Erlbaum Associates. Hillsdale.
- Evans, W. J. (2000). *Construct Validity of the Attitudes About Reality Scale*. *Psychological Reports*, 86(3, Pt1), 738-744.
- Fischhoff, B. (1991). *Value Elicitation: Is there Anything in There?* *American Psychologist*, 46, 835-847.
- Fraenkel, J.R and Wallen, N.E. (2009). *How to Design and Evaluate Research in Education*. (7thed). New York. The McGraw-Hill Companies.
- Fraenkel, J.R and Wallen, N.E. and Hyun, H.H. (2012). *How to Design and Evaluate Research in Education*. (8thed). New York. The McGraw-Hill Companies.
- Frankena, William K.; Raybeck, Nathan; Burbules, Nicholas (2002). *Philosophy of Education*. In Guthrie, James W. *Encyclopedia of Education*, 2nd edition. New York. Macmillan Reference.
- Friedman, Milton. (1962). *Role of Government in Education*. New Brunswick, New Jersey, United States of America. Rutgers University Press.
- Fowler, F.J. (2002). *Survey Research Methods* (3rded). Thousand Oaks, CA: Sage Publication.
- Furqon. (2004). *Masih Perlukah Ujian Nasional? Is the National Exam Still Needed?*. *Pikiran Rakyat*. Retrieved December 6, 2007, from <http://www.pikiran-rakyat.com/cetak/1204/23/0804.htm>
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational Research: An Introduction* (6thed.). United States of America. White Plains, NY: Longman Publishers.
- Gipps, C. (1999) *Socio-Cultural Aspects of Assessment*. *Review of Research in Education*, 24, 355-392.
- Grayling, A. C. (1999). *Philosophy: A Guide Through the Subject, Vol. 1*. United States of America. Oxford University Press.
- Green, K.E., & Stager, S.F. (1986). *Measuring Attitude of Teachers towards Testing*. *Measurement and Evaluation in Counseling and Development*.
- Grima, G. (2003). *School-Based Assessment : Implementation Issues and Practices*. Paper. (August). Cape Town South Africa. University of Malta.

- Haladyna, T., Nolen., S.B., & Haas, N.S. (1991). *Raising Standardized Achievement Test Scores and the Origins of the Score Pollution*. Educational Researcher, 20 (5), 2-7.
- Hamlyn, D.W. (1995). *Epistemology, History of the Oxford Companion to Philosophy*. Oxford. Oxford University Press.
- Herman, J.L. (1989). *Priorities of Educational Testing and Evaluation: The Testimony of the CRESST National Faculty* (CSE Tech. Rep. No. 304). Los Angeles, CA: University of California, Center for the Study of Evaluation.
- Herman, J., & Golan, S. (1991). *Effect of Standardized Test on Teachers and Learning – Another Look*. CSE Technical Report. Los Angeles: Center for the Study of evaluation.
- Hofer, B. K., & Pintrich, P. R. (1997). *The Development of Epistemological Theories: Beliefs about Knowledge and Knowing and Their Relation to Learning*. Review of Educational Research, 67, 88-140
- Hofer, B. (2001). *Personal Epistemology Research: Implications for learning and teaching*. Educational Psychology Review, 13, 353–384.
- Howard, B. C., McGee, S., Schwartz, N. H., & Purcell, S. (2000). *The Experience of Constructivism: Transforming Teacher Epistemology*. Journal of Research on Computing in Education.
- Humphreys, P. (1989). *The Chance of Explanation: Causal Explanation in the Social, Medical, and Physical Science*. Princeton. Princeton University Press.
- Hughes, A. (2003). *Testing for Language Teachers (2nd Ed)*. Cambridge. Cambridge University Press.
- Johnson, Burke & Christensen, L. (2012). *Educational Research. Quantitative, Qualitative, and Mixed Approaches*. 4th Edition. California. SAGE Publications. Inc.
- Jackson, Dionne Bennet. (2010). *The Impact of Science Teachers' Epistemological Beliefs on Authentic Inquiry: A Multiple-Case Study*. Theses. Baylor University.
- Johnson, R.B., & Onwuegbuzie, A. J. (2004). *Mixed Method Research: A Research Paradigm whose Time has Come*. Educational Researcher, 33, 14 -26.
- Jones, B.D., & Egle, R.J. (2004). *Is Testing the Right Direction? Administrators share their thoughts*. ERS Spectrum, 22(3), 16-25.

*KEPUTUSAN PRESIDEN REPUBLIK INDONESIA Nomor 41 Tahun 1987 Tentang
Pembagian Wilayah Republik Indonesia Menjadi 3 (Tiga Wilayah Waktu)
Presiden Republik Indonesia*

- Kneller, George (1964). *Introduction to the Philosophy of Education*. New York, NY: John Wiley & Sons. p. 93.
- Kok-devries, M. (2011). *Standardized Testing and the Impact on Classroom Instruction*, (July), 1–11. Omaha. University of Nebraska.
- Kumar, Ranjit. (2005). *Research Methodology: A Step by Step Guide for Beginners*. 2nd Ed. London. SAGE Publication.
- Kurt, Fatma. (2009). *Investigating Students' Epistemological Beliefs Through Gender, Grade Level, And Fields of the Study*. The Middle East Technical University.
- Levy, Paul. S & Lemeshow, S. (2008). *Sampling of Populations: Methods and Applications*. 4th Ed. United States of America. A John Wiley & Sons, Inc., Publication.
- Lichtenstein, S., & Slovic, P. (1971). *Reversals of Preference Between Bids and Choices in Gambling Decisions*. *Journal of Experimental Psychology* 89, 46-55.
- Madaus, George F, et. al. (1983). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. United States. Kluwer Academic Pub.
- Madaus, G. F. (1988). *The Distortion of Teaching and Testing: High-Stakes Testing and Instruction*. *Peabody Journal of Education*, 65 (3), 29 – 46
- Madaus, G. (1988). *The Influence of Testing on the Curriculum*. In L. Tanner (Ed.), *The Politics of Reforming School Administration*. London: Farmer Press.
- Mann, Prem S. (1995). *Introductory Statistics* (2nd ed.). Canada. John Wiley & Son.
- Magee, R. G., Jones, B. D., & Tech, V. (2012). *An Instrument to Assess Beliefs about Standardized Testing: Measuring the Influence of Epistemology on the Endorsement of Standardized Testing*, 12(0311), 71–82.
- Maynard, M. (1994). *Methods, Practice, and Epistemology: the Debate about Feminism and Research 'Researching Women's Lives from Feminist Perspective*. London. Taylor and Francis.
- Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). *When Less is More: Counter Factual Thinking and Satisfaction Among Olympic Medalists*. *Journal of Personality and Social Psychology*, 69,603410.

- Mohamad Najib Abdul Ghafar. (2003). *Reka Bentuk Tinjauan Soal Selidik Pendidikan*. Skudai. Johor Bahru. UTM Penerbit.
- Mohamad Najib Abdul Ghafar. (2009). *Prinsip Asas Penilaian Program Pendidikan*. Skudai. Johor Bahru. UTM Penerbit.
- Mohamad Najib Abdul Ghafar. (2011). *Pembinaan dan Analisis Ujian Bilik Darjah*. 2nd Ed. Skudai. Johor Bahru. UTM Penerbit.
- Mohamed Najib Abdul Ghafar. (2015). *Methodology in Behavioural Research*. Skudai. Penerbit UTM Press.
- Montgomery, H. (1983). *Decision Rules and the Search for A Dominance Structure: Toward A Process Model of Decision Making*. In P. C. Humphreys, Amsterdam: North-Holland.
- Morse, J.M. (1991). *Approaches to Qualitative-Quantitative methodological Triangulation*. Nursing Research, 40, 120 – 123.
- Moser, C.A & Kalton, G. (1989) *Survey Methods in Social Investigation*. 2ndEd. Boston USA. Dartmouth Publishing.
- Nolan, S.B., Haladyna., T.M & Haas, N.S. (1992). *Uses and Abuses of Achievement Test Scores*. Educational Measurement: Issues and Practice II(2), 9-15.
- Oakes, J., & Lipton, M. (2007). *Teaching to Change the World* (3rd ed.). Boston: McGraw-Hill. Methods of data Collection.
- Olafson, L. & Schraw, G. (2006). *Teachers' Beliefs and Practices Within and across Domains*. International Journal of Educational Research.45, 71-84.
- Onwuegbuzie, A.J., & Collins, K.M.T. (2007). *A Typology of Mixed Method Sampling Design in Social Science Research*. The Qualitative Report, 12 (2), 281-316. (ERIC Document Reproduction Service No. EJ800183)
- Overton, Terry. (2012). *Assessing Learners with Special Needs. An Applied Approach*. 7th Edition. New Jersey United States of America. Pearson Education, Inc.
- Parahoo, K. (1997). *Nursing Research: Principles, Process, and Issues*. United States of America. Palgrave Macmillan
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). *Behavioral Decision Research: A Constructive Processing Perspective*. Annual Review of Psychology, 43, 87-131.
- Pedulla, J. P. (2003). *State-Mandated Testing – What do Teachers Think?* Educational Leadership, 61 (3), 42 – 46.

- Perry, W. G. Jr. (1970). *Forms of Ethical and Intellectual Development in the College Years: A scheme*. New York: Holt, Rinehart & Winston.
- Philibber, S.G, Schwab, M.R & Sam Loss, G. (1980). *Social Research*. Adelaide Australia. F.E. Peacock Publishers.
- Piaget, J. (1970). *Genetic Epistemology*. New York: Columbia University Press.
- Polit, D. F. et al. (2001) *Essentials of Nursing Research: Methods, Appraisal and Utilisation* (5thed). Philadelphia. PA: Lippincott.
- Polit, D. F & Hungler, B. P. 1999. *Nursing Research: Principles and Methods*. 6thed. Philadelphia. Lippincott.
- Portal, M. (2003). *Classroom Assessment and its Consequences*. Proceedings of The Second International Conference of the Association of Commonwealth Examinations and Accreditation Bodies. Malta: MATSEC Examinations Board.
- Pratt, D. D. (1998). *Five Perspectives on Teaching in Adult and Higher Education*. Malabar, FL: Krieger Publishing Company.
- Queensland Studies Authority. (2007). *School-Based Assessment in Year 11-12*. Retrieved on 27 August 2007 from <http://www.qsa.qld.edu.au/yrs1112/assessment/docs/school-based-assess-fact-sheet.pdf>
- Quinton, Anthony. (1995). *The Ethics of Philosophical Practice*. United States of America. Oxford University Press.
- Raffan, J. (2001). *School-Based Assessment: Principles and Practice*. Proceedings of The First International Conference of the Association of Commonwealth Examinations and Accreditation Bodies. Reduit: Mauritius Examinations Syndicate.
- Rand, Ayn. (1982). *Philosophy: Who Needs It*. Inc. 630 third avenue New York. United States of America. Bob-Merril Co.
- Rohaya, T., MohdZaki, K., Hamimah, A. N., & Adibah A. L. (2014). From Principle to Practice: Assessment for Learning in Malaysian School-based assessment Classroom. *International Journal of Social Sciences & Education*, 4(4), 850 – 857.
- Romberg, T.A., Zarinnia, A., & Williams, S.R. (1989). *The Influence of Mandated Testing on Mathematics Instruction: Grade 8 teachers' perceptions*

- (*Monograph*). Madison, WI: University of Wisconsin, National Center for Research in Mathematical Science Education.
- Roth, W. M., & Roychoudhury, A. (1994). Physics Students' Epistemologies and Views about Knowing and Learning. *Journal of Research in Science Teaching*, 31(1), 5-30.
- Ryan, A. G., & Aikenhead, G. S. (1992). *Students' Preconceptions about the Epistemology of Science*. Science Education.
- Salmon, M.H. (2007). *Introduction to Logic and Critical Thinking*. Belmont, CA: Thompson Higher Education.
- Sapsford, Roger & Jupp, Victor. (1996). *Data Collection and Analysis*. 2nd Edition. SAGE Publication. London.
- Sax, Gilbert. (1980). *Principles of Educational and Psychological Measurement and Evaluation*. 4th edition. Wadsworth Publishing Company. Belmont, California. USA.
- Sax, Gilbert. (1997). *Principles of Educational and Psychological Measurement and Evaluation*. 2nd edition. Wadsworth Publishing Company. Belmont, California. USA.
- Selltize, Claire, et.al. (1965). *Research Methods in Social Relations*. Methuen & Co Ltd. United States of America.
- Schommer, M. (1990). *Effects of Beliefs about the Nature of Knowledge on Comprehension*. *Journal of Educational Psychology*, 82, 498-504.
- Schommer, M. (1994). *Synthesizing Epistemological Belief Research: Tentative Understandings and Provocative Confusions*. *Educational Psychology Review*, 6(4), 293-319
- Schommer-Aikin, M. (2002). *An Evolving Theoretical Framework for an Epistemological Belief System*. In B.K. Hofer & P.R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 103-118). Mahwah, N.J.: L. Erlbaum.
- Shawer, S.F., Gilmore, D., & Banks-Joseph, S. (2008). *Student Cognitive and Affective Development in the Context of Classroom-level Curriculum Development*. *Journal of the Scholarship of Teaching and Learning*.
- Shawer, S.F., Gilmore, D., & Banks-Joseph, S. (2009). *Learner-Driven EFL Curriculum Development at the Classroom Level*. *International Journal of Teaching and Learning in Higher Education*.

- Shawer, S.F. (2012). *Standardized Assessment and Test Construction without Anguish. The Complete Step-by-Step Guide to Test Design, Administration, Scoring, Analysis, and Interpretation*. New York. Nova Science Publishers, Inc.
- Shepard, L.A. (1990). *Inflated Test Score Gains: Is it old norms or teaching the test?* Educational Measurement: Issues and Practices, 9(3), 15-22.
- Slovic, P., Griffin, D., & Tversky, A. (1990). *Compatibility Effects in Judgement in Choice*. In R. M. Hogarth (Ed.), *Insight in decision Making: A Tribute to Hillel J, Einhorn* (pp 5 - 27). Chicago. University of Chicago Press.
- Slovic, P. (1995). *The Construction of Preference*. American Psychologist, 50, 364-371.
- Smith, M.L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). *The Role of Testing in Elementary Schools (Monograph)*. Tempe, AZ: Arizona State University, Center for Research on Evaluation, Standards, and Student Testing (Grant No. OERI-G-86-0003).
- Schraw, G., & Olafson, L. J. (2002). *Teachers' Epistemological Worldviews and Educational Practices*. Issues in Education, 8, 99–148.
- Stake, R.E. (1988). *The Effect of Reforms in Assessment in the USA*. Paper presented at the annual meeting of the British Educational Association at the University of East Anglia.
- Stopher, P. (2012). *Collecting, Managing, and Assessing Data Using Sample Surveys*. United Kingdom. Cambridge University Press.
- Struening, E. & Guttentag, M. (1975). *Handbook of Evaluation Research*. Beverly Hills. Sage Publication. (VL1).
- Teichmann, Jenny and Evans, Katherine C. (1999). *Philosophy: A Beginner's Guide*. United States of America. Wiley.
- Tversky, A., Sattath, S., & Slovic, P. (1988). *Contingent Weighting in Judgment and Choice*. Psychological Review, 95, 371-384
- Tuah, Adi Badiozaman (2006). *Improving the Quality of Primary Education in Malaysia through Curriculum Innovation: Some Current Issues on Assessment of Students' Performance and Achievement*. Proceedings 3rd International Conference on Measurement and Evaluation in Education (ICMEE), pp. 16-26. Penang: University Science of Malaysia.

- UNESCO. (2005). *Country Report at the UNESCO Seminar on "EFA Implementation: Teacher and Resource Management in the Context of Decentralization"*. Hyderabad India. Administrative Staff College of India.
- UUD SISDIKNAS No. 20. (2003) *Undang-Undang Republik Indonesia No. 20 Tahun 2003. Sistem Pendidikan Nasional*.
- Walter, Benjamin A. (2009). *Epistemological Belief: Differences Among Educators*. Thesis. Wichita State University.
- Weisberg, D.M., Bowen, B.D. (1977). *An Introduction to Survey Research and Data Analysis*. San Francisco. W. H. Freeman and Company.
- Weisberg, D.M., Krosnick, J.A., Bowen, B.D. (1996). *An Introduction to Survey Research, Polling, and Data Analysis*. (3rd Edition). United States of America. Sage Publication, Inc.
- Wood, L.G. & Judith, Haber. (1998). *Nursing Research: Methods, Critical Appraisal, and Utilization 4th edition*. United States of America. Mosby-Year Book.
- Yin, Robert. K. (1989). *Case Study Research. Design and Method*. United States of America. Sage Publication.
- Yong, Hwa Tee, Sam, Lim Chap, (2008). *Implementing School Based Assessment: The Mathematical Thinking Assessment (MATA) Framework*. Book of Seminar of Pedagogical Innovation. Universiti Teknologi Mara. Sarawak and Universiti Sains Malaysia, Penang.