A Confirmatory Factor Analysis of Classroom Assessment Practises Scale in a Malaysian Context

Nor Hasnida Che Md Ghazali, Norfishah Mat Rabi, Nurulhuda Md Hassan, Norwaliza Abdul Wahab

Faculty of Human Development, Universiti Pendidikan Sultan Idris, Malaysia
Email: hasnida@fpm.upsi.edu.my

Abstract
Classroom assessment practices are an important element in improving the teaching and learning process. Previous research proposed that classroom assessment practices among teachers should be categorized into five dimensions which are assessment planning, assessment item preparation, assessment administration and scoring, reporting of scores and grading and the last is assessment data utilization and evaluation. This study is aimed to validate a Classroom Assessment Practises (CAP) instrument which consists of 56 items altogether. This study is a quantitative approach which involves the collection of data using questionnaire. The sample is 320 secondary school teachers teaching Islamic education in one of the states in Malaysia. Exploratory factor analysis and confirmatory factor analysis were used to examine CAP using AMOS 16.0. Some constructs revealed acceptable internal consistency reliability. A good model fit was found for the measurement model using several fit index tests like CMINDF, GFI, CFI, TLI and RMSEA. The findings showed that all fit indices criteria were fulfilled. It also showed that acceptable construct reliability and variance extracted value were obtained. This validated instrument could then be used for real study in various context. Teachers could also use it as a self-assessment tools to look for their strengths and weaknesses.
Keywords: Classroom Assessment Practises, Confirmatory Factor Analysis, Exploratory Factor Analysis, Islamic Education Teachers

Introduction
Globalization and internationalization are the two challenges that demand a dynamic transformation in the education system in Malaysia (Ganapathy & Kaur, 2014). Besides fulfilling the 21st century global needs, responding to such challenges is also important to ensure the achievement of advanced nation status by 2020. This notion is supported by Nooraini and Azmi (2014) who claimed that the development of thinking skills among students has been among the objectives of education system in order to help Malaysia in achieving Vision 2020. The transformation in the education system in Malaysia is manifested in the National Education Blueprint (2013-2015) whereby articulation of the specific skills and attributes including higher order thinking skills has been refined. Such refinement is vital as a

preparation in producing knowledgeable students with real life skills so that they can compete in the global labor market. In other words, the new reform in learning systems emphasizes the importance of cultivating student's higher level thinking abilities (Kim, 2005). And, the new reform has to include all the three most important elements in teaching and learning process which are curriculum, teaching strategy and assessment.

There are two dominant methods in assessing students if we were to help students learn well (Shepard, 2000). The first one refers to the content and the nature of the assessment meanwhile, the second method is the collection and the use of assessment information to be part of a continuous learning process. However, by means of the assessment approach, there are two ways of assessing students used by teachers. They are one-off and on-going approach. The one-off approach is an approach that requires the implementation of the assessment process on student performance in a separate test during the teaching and facilitating process. On-going approach is a valuation approach that is administered during the teaching and facilitating process. It is also considered as part of the teaching and facilitating process, and is carried out continuously.

By definition, classroom assessment is an assessment that is conducted in the classroom by the teachers (Black and Wiliam, 2004). The materials used in the classrooms could come from the teachers or any external agents. The assessment methods could be traditional written examinations, oral examinations, open-book examinations, resources used from the internet, tests or quizzes. Another way of assessing students could be communication with students, informal assessments, content area inventories, classroom works, rubrics and scoring guide, writing journals or using checklists (Suzieleez et al., 2009). As classroom assessment is meant to be conducted in classroom only, strategies which could be used in classroom can be applied. Moreover, classroom assessment gives teachers the opportunity to take immediate and appropriate action to improve planning, teaching methods and approaches (Black & William, 1998). McMillan (2007) expresses classroom assessments as a process that advocates and enhances student learning, and is not just a document about what students know, understand and do. Classroom assessment is also described as a process that encompasses a wide scope of speech and observation records up to the standard test stage (Azizi, 2010). It needs to be done continuously by whatever means in helping pupils to improve learning in each teaching unit. Classroom assessment could either be formative or summative depending on its type. If the assessment is implemented when teachers and students provides feedback to adjust an ongoing teaching and learning to improve students' achievement (Popham, 2008),it is called formative and if it is measuring students' level of attainment in the form of monitored examination (Rayment, 2006), it is called summative.

When assessing their students in a classroom, teachers have to focus on the four components. First is to look at why do they assess?, followed by what are they going to assess based on the syllabus?, how are they going to assess?, and lastly is how to interpret the results (Azman Wan Chik, 1982). According to Chappius et al. (2012), teachers have to have good knowledge and literacy skills in order to conduct classroom assessment. Teachers have to know how to compile the data and produce the assessment results effectively to improve students' achievement. In addition, the important thing for a quality class assessments is to design the assessment to meet the needs of certain data users, based on the assessment of the concrete and appropriate achievement targets, to determine precisely the achievement of the

students, to produce evaluation results that effectively communicate to users and involving student participation in self assessment, goal setting, monitoring, reflection, and sharing of learning among students.

Suah and Ong (2012) conducted a study on 406 teachers stated that teachers might give importance to the kind of assessment that is suitable for teaching on a given day. In addition, most teachers have rarely used the test-setting tables to build test items. The main source of reference used by most teachers when constructing test questions is textbooks. This is followed by the use of reference books and questions from public examinations. In addition, it is found that teachers are more committed to the various forms and homework assignments to evaluate pupils' learning. Test questions are usually at a low level of understanding, which is based on Bloom's Taxonomy followed by application level, knowledge, analysis, assessment and synthesis.

A limited teacher's knowledge of a subject is found to cause teachers to focus more on certain concepts, and not be able to establish relationships between facts, concepts, structures and practices. The weak teacher's knowledge of subjects creates limitations in applying various methods for understanding pupils (Walshaw, 2012). As such, the teacher's knowledge of various classroom evaluation practices is very important. The findings of the study by Mohd Fadhli Ahmad (2010) show the level of understanding on the practice of assessment among teachers at a high level. This high level of mastery is contributed by understanding and mastery of basic concepts of assessment, general principles of testing, planning of classroom assessments, test determination tables and item difficulty levels. According to his research, there are five principles to be considered in ensuring high-quality assessments of legality, reliability, objectivity, administration and interpretation. The findings of the study by Mohd Fadhli Ahmad (2010) show the level of understanding on the practice of assessment among teachers is at a high level. This level of mastery is contributed by understanding and mastery of basic concepts of assessment, general principles of testing, planning of classroom assessments, test determination tables and item difficulty levels. According to his research, there are five principles to be considered in ensuring high-quality assessments which are validity, reliability, objectivity, administration and interpretation.

Next, feedback is an important component in classroom assessment process (Nicol & Macfarlane-Dick, 2010). Feedback refers to the information that describes the current situation of pupils in terms of learning and performance. Hence, the primary purpose of feedback is to develop students in achieving personal goals and to help them to compare or evaluate performance independently. The Black & William (1998) study also noted that effective feedback led to increased student learning processes. However, Black & William (1998) found that the meaningful assessment process of classroom assessments was difficult to explore if teachers were still using traditional forms of assessment. Nisbet and Warren (2000) also conducted a study on 398 respondents indicated that teachers believe that classroom assessments play an important role in informing teachers and students about academic achievement as well as for accountability purposes. The results of his study found that the use of assessments aimed at informing teachers is less practiced when the level of student learning progressively increases.
Furthermore, the success of the implementation of HOT skills in education also depends on the assessment practices carried out by teachers. In Malaysia, students' ability to solve HOT

questions can be fostered through the implementation of the assessment framework introduced by the Ministry of Education (Abdullah et al., 2017). Therefore, teachers are not only required to be competent in the aspects of curriculum and pedagogy of HOT skills, but they also need to possess the knowledge of assessment in order to ensure the successful implementation of HOT skills. The change in PPPM 2013-2015 i.e. upgrading the quality of education to international standard through the implementation of HOTS practices requires Malaysian Examination Board to change the assessment and examination systems. In response to this, assessments carried out in classrooms should emphasize the aspects of HOT skills (Ganapathy & Kaur, 2014). Apart from the application of HOT skills, Zhang and Judith (2003) argue that an assessment should also ensure the continuity of students' learning. According to Nenty et al. (2007), an effective assessment has two main characteristics; allows the recognition of students' level of learning and acknowledges the effectiveness of teachers' teaching activities. Above from that, Zhang and Judith (2003). Abdullah et al. (2017) found that teachers have weak proficiency in the aspect of HOTS assessment. This situation resulted in their inability to shift their traditional assessment practices which employ exam and teacher-centered orientation (BPK, 2012; Caroline & Abdul Said, 2014).

Morshidi Sirat et al. (2008) found that lack of thinking skills was one of the factors that decrease the employability among Malaysian students who qualify for tertiary education and graduate from the system. Due to this situation, The Ministry of Education has considered this aspect by requiring the integration of HOTS elements in the questioning format for the national examinations and school-based assessments (Ganapathy & Kaur, 2014). Such integration is done in order to ensure 40% of questions in the Upper Primary National Assessment (UPSR) and 50% in the Malaysian Education Certificate (SPM) comprise higher order thinking questions by 2016.

Makeleni and Sethusa (2014) claimed that teachers' knowledge and understanding of assessment practices in education determine the extent to which an assessment is effective in fostering the application of HOT skills among students. Besides enhancing students' learning, such knowledge and understanding are also important in improving teachers' teaching methods (Serdyukova, 2015) as well as helping students to develop self-assessment skills (Anthony & Walshaw, 2009). Abdullah et al. (2017) assert that incomprehensive and inaccurate exposure to the knowledge of assessment led to the weakness in the aspect of assessment among teachers. Therefore, in order to improve the quality of classroom and school assessment, Rasidayanty (2014) stated that proper program and training should be given to teachers so that they can enhance their knowledge of assessment. This is due to the reason that specific assessment strategy only appropriate for assessing specific learning objective (Hammond, 2006). Consequently, teachers need to be skillful in every type of assessment as a means to ensure the match between the assessment strategy and the thinking skills that need to be assessed (Abdullah et al., 2017).

The model used to support this study is a model developed by Chris Rust and his colleagues from the Oxford Centre for Staff and Learning Development named The Social Constructivist Assessment Process Model (SCAP Model) (Rust *et al.,* 2005). This model is created as an initiative to lead assessment practises in classroom based on a social constructivist approach. By means of social constructivist approach, knowledge is seen as mutually-constructed or co-constructed with the combination of someone's prior knowledge with new knowledge

influences by past and present social context (Brooks, 2002). The overall model consists of two sub-models with parallel ongoing cycles of four different stages for each model. One model is applied for staff and the other one is for students. These two models are then combined to form a dynamic system of the SCAP Model. A very important requirement for a constructive assessment process is to have well-defined explicit assessment criteria. Next, students and teachers have to be actively-engaged with the criteria for the students to have a meaningful understanding of the assessment to improve their work, such as students marking their own work, peer marking, teachers having peer-discussions on criteria, marking process or moderation process. It is believed that it is better for the teachers to create and make decisions on the assessment criteria rather than be given finished criteria by the course leader; however, for students, it is the other way round. Following active-engagement with the criteria is active-engagement with the feedback by the students. Constructive feedback lies at the heart of assessment processes, so active engagement in it is very important. Feedback does have an effect on students' learning as long as it is practised in a correct manner and students are actively engaged in it. Teachers have to train their students on how to deal with feedback to get the benefit out of it. But, how? Firstly, responding to teachers' feedback, students are encouraged to have peer discussions by referring to the checklists and come out with suggestions. Secondly is by providing comments without assigning students with any grades or marks. Similarly to this, it could also be a reflective self-assessment made by the students before the teachers return back the students' work. Thirdly is to give a general generic feedback to the whole class as soon as their work is completed and not to provide them with detailed individual feedback. Effective feedback gives way to staff to design assessment and develop explicit assessment criteria. Then, the criteria would have to be clearly explained to students by the staff and the process continues in the form of a cycle.
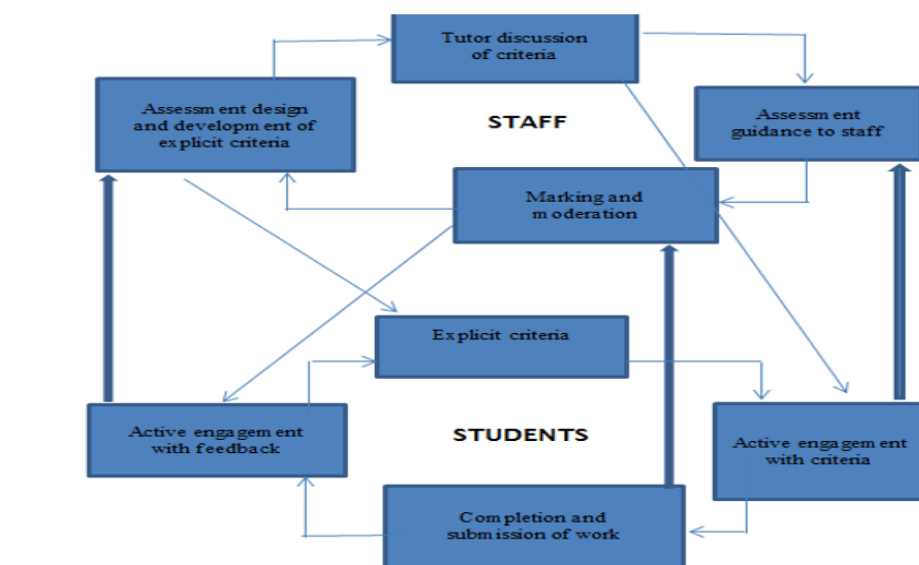
Objectives of the Study

Until now, in the Malaysian context, there is no thorough study regarding the development of instrument on CAP in schools. Hence, the goal of this present study is to validate the existing instrument developed by Gonzales (2012). This psychometrically sound (valid and reliable) instrument can be very useful for researchers and educators interested in determining the implementation of CAP in schools.

Methodology

This study is a quantitative approach which involves the collection of data using questionnaire. The sample involved were 320 secondary school teachers teaching Islamic education subject from all over Malacca, a state in Malaysia. Simple random technique is used in choosing the sample of the study. Most of the teachers are female (79.9%), and only 20.1% of the sample are male teachers. 81.9% of them are Malays, 27% of them are Chinese, 6.7% are Indians, and the balance are from other ethnic groups. The instrument used was developed by Richard DLC Gonzales (2012), which consisted of 56 items. A 5-point Likert scale ranging from 1 (never) to 5 (always) is used. The instrument was administered by the researcher and the respondents were given ample time to answer the survey. The representative of each school collected the answered survey and returned them to the researcher. The difference in research setting has become the justification in testing the CAP among the Islamic education teachers in Malaysia. Prior to conducting the real data collection, the researcher conducted a pilot study with 100 respondents. Comments and recommendations during this pilot study have been considered by the researcher in order to improve the instrument. The variability in cultural setting as well as checking the validity and reliability of CAP are required in this context of study.

Findings and Results

In order to validate the underlying hypothesized factor structure of CAP instrument, exploratory factor analyses and confirmatory factor analyses were used.

a. Exploratory Factor Analysis

EFA involves several processes to ensure that the factor consists of suitable items (Tabachnick & Fidell, 2001). According to Singer and Willet (2003), it has to be applied to the data before fitting with the models. This technique is used to reduce the number of items in a particular construct in order to improve the variance and reliability values. It also helps in identifying the dimensions which might be included (Netemeyer *et al*., 2003). To fix the appropriate number of factors for the scale, the researcher used the eigenvalue rule and scree plot test. The Barlett Test of Sphericity was significant (p=0.00). Items with factor loading of more than 0.40 were retained since this would likely increase the reliability of the scale (Costello & Osborne, 2005). Only 21 items were considered in the CAP at the final stage of the EFA process. The results of the EFA revealed that CAP scale yielded three factors for the respondent of the study instead of five factors originally.

b. Reliability

The value of Cronbach alpha obtained should be at least 0.70 for any research using the survey method (Hair *et al*., 2010). The reliability coefficient of CAP is 0.878 and each subscale of CAP has reliability value ranging from 0.747 to 0.914 (Table 1). Thus, CAP has an acceptable value of internal consistency for this study.

Table 1
Overall Cronbach Alpha Value for the Assessment Practises Constructs

Construct	No of items	Item	Overall Cronbach Alpha Value
Item preparation	7	c20, c21, c22, c23, c29, c39, c41	0.747
Administration and Scoring	8	c35, c36, c37, c38, c44, c46, c48, c51	0.914
Utilization and evaluation	6	c49, c50, c52, c54, c55, c56	0.887

c. Confirmatory Factor Analysis

This theory driven technique called CFA is technique used to confirm a measurement model (Byrne, 2010). It is achieved by determining the goodness of fit between the model and the sample data. This technique estimates a population covariance matrix from the observed covariance matrix from the hypothesized model and it minimizes the difference between the two matrices. In addition, it is suitable when the researcher has had some knowledge from the empirical research (Arbuckle, 2010). In order to achieve model fit, fit statistics tests like traditional chi-square test, the relative chi-square (CMINDF: the chi-square/degree of freedom), Tucker Lewis Index (TLI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI) and Root Mean Square of Error Approximation (RMSEA) were chosen (Hair *et al.*, 2010). The acceptable criterion for traditional chi-square is shown by non-significant result. The relative chi-square (CMINDF) must be between 1 and 5 in order to achieve the fitness of the model. The TLI, CFI, and GFI values should be in the range of 0 to 1. However, the RMSEA value should fall below 0.08 to indicate an acceptable fit to the data (Schumacker & Lomax, 2004).

Assessment practices consists of five major constructs with **56 items** altogether. Then, the reliability is checked as shown in Table 4.11. The AMOS Output is showed in Table 2.

Table 2
AMOS Output for the hypothesized model for summary statistics

Computation of degrees of freedom number	
Number of distinct sample moments	231
Number of distinct parameters to be estimated	43
Df	188
Result	
Chi-square	1884.852
Df	188
Probability level	0.000

In reviewing the goodness-of-fit statistics, it showed that $X^2/df=10.033$, CFI=0.662, IFI=0.664, TLI=0.623, GFI=0.667 and RMSEA=0.168. This measurement model provided a poor fit (Figure 1). Four items are deleted due to low factor loading value. Then, we look at the modification index values. Few items were deleted. Again, the model is still not fit.
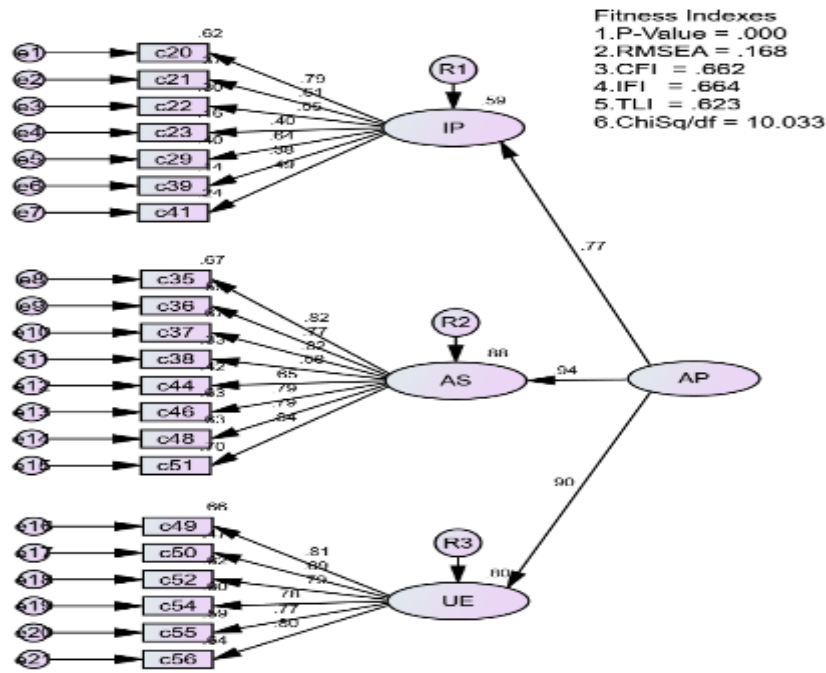
Figure 1. Measurement Model

Finally, 12 items were deleted. It showed that X^2/df = 6.741, CFI = 0.916, IFI = 0.917, TLI = 0.879, GFI = 0.901 and RMSEA = 0.134.
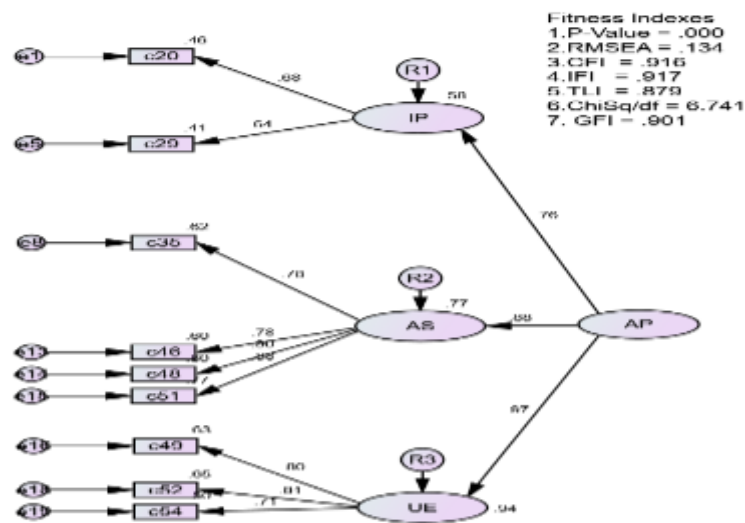

Figure 2. The finalized measurement model

So, it was a three-factor model with 9 items (Figure 2). Byrne (2010) asserted that assessment of model adequacy should be based on theoretical, statistical and practical considerations. Looking at Table 3 the value of skewness indicated that every item was approximately normally distributed. The value of kurtosis was also acceptable and the value of multivariate kurtosis which was lower than 50.0 indicated multivariate normality distribution of the data set. Finally, the issues of uni-dimensionality, validity and reliability were addressed in Table 4.

Table 3
Assessment of normality

Variable	min	max	skew	c.r.	kurtosis	c.r.
c49	3.000	5.000	-.003	-.025	-.333	-1.217
c52	2.000	5.000	-.761	-5.559	1.539	5.619
c54	2.000	5.000	-.524	-3.828	.443	1.618
c51	3.000	5.000	-.121	-.883	-.542	-1.978
c35	3.000	5.000	-.011	-.082	-.115	-.419
c46	2.000	5.000	-.275	-2.008	.349	1.275
c48	2.000	5.000	-.189	-1.382	.084	.308
c20	3.000	5.000	-.094	-.690	-.739	-2.697
c29	3.000	5.000	.033	.241	-.414	-1.512
Multivariate					28.987	18.426

Table 4
The CFA results reporting for the measurement model

2nd order factor	1st order factor	Standardized Factor Loading			
AP	IP	0.762			
	AS	0.880			
	UE	0.972			
1st order factor	**Item**	**Standardized Factor loading (>0.5)**	**CR Alpha (>0.7)**	**CR (>0.6)**	**AVE (>0.5)**
Item Preparation (IP)	c20	0.680	0.715	0.606	0.435
	c29	0.638			
Administration and Scoring (AS)	c35	0.785	0.896	0.902	0.698
	c46	0.776			
	c48	0.896			
	c51	0.879			
Utilization and Evaluation (UE)	c49	0.795	0.815	0.814	0.594
	c52	0.807			
	c54	0.707			

Discussion

Recently, there is no instrument in Malay language which measures the implementation of CAP especially in the context of secondary school teachers teaching Islamic education subject. Therefore, this study is conducted to measure the psychometric properties of an existing instrument used to measure CAP. This act of validating the instrument to the context which we are interested at is very important as it will give the researcher confidence with the quality of items to be used in real studies for further research. The measurement model has yielded three factors which are item preparation, administration and scoring and utilization and evaluation. At the beginning of the study, the existing instrument consists of 5 constructs with

56 items altogether. After EFA has been conducted, it left with three constructs only. After EFA was conducted, the measurement model left with 21 items. After CFA was applied, 9 items are maintained. At this point, the instrument is valid and reliable. The EFA and CFA analyses provide evidence to the reliability and validity of CAP. This is shown by the construct reliability as well as variance-extracted value that exceeds the suggested value. Therefore, CAP can be replicated in future works with a wider sample range. This would direct to the betterment of the instrument, which could be used in measuring teachers' assessment practices.

Implications and Suggestions for Future Research
The results of the literature review and the informal interviews lead the researcher to review the assessment planning and item preparation and also the challenges that teachers have to face in implementing CAP. Furthermore, with this valid instrument, the researcher would be very confident in using this instrument in real study later in the Malaysian context. In order to obtain a better understanding of this adapted instrument, it would be advisable to collect data from primary schools in states other than Malacca. This is to test the validity of the study's model across different school samples and the extent to which these can be generalized. This validated instrument could also be used as a self-assessment tools for teachers. Teachers could then determine their strength and weaknesses. Teachers should be given flexibility in assessing their students in a summative way or in a standardized form because in order to improve education quality, assessments should be integrated in the teaching and learning process (Kelvin, 2007).

In future research, the researcher might use the variables developed in this study to look for the interrelationships between variables with other variables such as school improvement, for example. This is important as the interrelationships between variables reflect how effective the system are (Nor Hasnida, 2015). This instrument could also be used as a pre-test and post-test when training is conducted. The training could be formal but Slaalvik (2010) believes that stimuli such as praise, encouragement and demonstration can also increase the level of teacher practices towards the implementation of CAP. So, in this case, this instrument could also be used informally.

Conclusion
The output for this study has shown the evidence of the reliability and the validity of the CAP. The acceptable model fit has been achieved. However, it is better for the instrument to go through further validity and reliability test with larger sample size in different context. The analyses yielded evidence that CAP can be a useful scale to measure the classroom assessment practices among teachers, especially in the Malaysian context. And, this measurement model could also be a starting point for further research to form other structural model of interest.

Acknowledgment

References

Abdullah, A. H., Mokhtar, M., Ali, D. F., Tahir, l. M. & Kohar, U.H.A. (2017). Mathematics teachers' level of knowledge and practice on the implementation of higher-order thinking skills (hots). *Eurasia journal of mathematics science and technology education*, 13(1):3-17.

Anthony, G. dan Walshaw, M. (2009). *Effective Pedagogy in Mathematics.* Switzerland: International Academy of Education.

Arbuckle, J. L. and Wothke, W. (1999). *Amos 4.0 User's Guide*. Small Waters Corporation, United States of America.

Azman Wan Chik (1982). *Pengujian bahasa: Kes Bahasa Malaysia.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning, *Assessment in Education*, Vol. 5, 7-71.

Black, P. and Wiliam, D. (1998). *Inside the Black Box: Raising standards through classroom assessment*. London: School of Education, King's College.

Black, P. and Wiliam, D. (2004). 'Classroom Assessment Is Not (Necessarily) Formative Assessment (and Vice-versa)', *Yearbook of the National Society for the Study of Education*, 103, pp. 183-188.

Byrne, B. M. (2010). *Structural Equation Modelling with AMOS. Basic concepts, applications and programming.* Mahwah: New Jersey: Earlbum.

Caroline L. D. dan Abdul Said Ambotang. (2014). Profesionalisme Guru Novis dalam Pengurusan Pengetahuan, Kesediaan Mengajar dan Kemahiran Berfikir Aras Tinggi (KBAT) Terhadap Pelaksanaan Pengajaran. Seminar Kebangsaan Integriti Keluarga 2014. 11 December, Universiti Malaysia Sabah.

Chappuis, J., Stiggins, R. Chappuis, S., Arter, J. (2012). Classroom assessment for student learning: Doing it right-using it well. Assessment Training Institute: London.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7).

Ganapathy, M. & Kaur, S. (2014). ESL Students' Perceptions of the use of Higher Order Thinking Skills in English Language Writing. *Advances in Language and Literary Studies*, 5(5): 81-87.

Gonzales, R. & Fugan, C. (2012). Exploring the Conceptual and Psychometric Properties of Classroom Assessment, *The International Journal of Educational and Psychological Assessment,* 9(2).

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. & Tatham, R. L. (2010). *Multivariate Data Analysis.* 6th edn. New Jersey: Prentice-Hall International, Inc.

Hammond, L. D. (2006). Constructing 21st century teachers' education. *Journal of Teacher Education,* 57(1): 1-15.

Kim, Y. (2005). Cultivating reflective thinking: *The effects of a reflective thinking tool on learners' learning performance and metacognitive awareness in the context of on-line learning.* (Ph.D dissertation, The Pennsylvania State University). Retrieved from http://search.proquest.com/docview/305419245?accountid=28930

Makeleni, N. T. & Sethusha, M. J. (2014). The experiences of foundation phase teachers in implementing the curriculum. *Mediterranean Journal of Social Sciences*, 5(2), 103-109.

McMillan, J. H. (2007). Classroom assessment: Principles and practice for effective mathematics, teaching and assessing mathematics and factors that influence these beliefs. *Mathematics Teacher Education and Development,* 2, 34-47.

Mohd Fadhli Ahmad (2010). Amalan Pentaksiran Pengajaran dan Pembelajaran di Kolej Komuniti di Negeri Johor. Prosiding Seminar Transformasi Pendidikan Teknikal anjuran Center for Testing, Measurement & Appraisal (CeTMA), Universiti Utara Malaysia, dari 3 hingga 4 November 2010.

Morshidi, S., Pandian, A., Muniandy, B., Fazal M., Harshita A., Hazita A., Muhammad K., & Ratna R.A.R. (2008). The university curriculum and the employment of graduates. In Zuraidah Mohd Don (Ed.), *Enhancing the Quality of Higher Education through Research: Shaping Future Policy* (pp.939). Kuala Lumpur: Ministry of Higher Education Malaysia, (MoHE).

Nenty, H. J., Adedoyin, O. O., Odili, J. N. & Major, T. E. (2007). Primary teacher's perceptions of classroom assessment practices as means of providing quality primary/basic education by Botswana and Nigeria. *Educational Research and Review,* 2 (4), 74-81.

Nicol, D. J. & Macfarlene-Dick, D. (2006) 'Formative assessment and self-regulated learning: A model and seven principles of good feedback practice', *Studies in Higher Education*, 31(2), 199-218.

Nisbet, S., & Warren, E. (2000). *Primary school teachers' beliefs relating to occurrence.* Oxford: Oxford University Press.

Nooraini Othman & Khairul Azmi Mohamad (2014). Thinking skill education and transformationalprogress in Malaysia. *International Education Studies,* 7(4), 27-32.

Nor Hasnida, CMG. (2015). An evaluation of the Implementation of the SBA System in Malaysia. (Doctoral Dissertation). Available at: https://eprints.soton.ac.uk/381724/

Popham, W. J. (2008). *Transformative Assessment*. Alexandria, Virginia: ASCD.

Rasidayanty Saion (2014). *Persepsi Guru-Guru Kemahiran Hidup Bersepadu Terhadap Literasi Pentaksiran Dalam Pelaksanaan Pentaksiran Berasaskan Sekolah.* Master of Technical Education and Vocational, University Tun Hussein Onn Malaysia.

Rayment, T. (2006). *101 Essential Lists on Assessment.* Kuala Lumpur: Institut Terjemahan Negara Malaysia Bhd.

Rust, C., O Donnovan, B. & Price, M. (2005) 'A social constructivist assessment process model: how the research literature shows us this could be best practice', *Assessment & Evaluation in Higher Education,* 30(3), 231-240.

Schumacker, R.E. & Lomax, R. G. (2004). A beginner's guide to structural equation modeling. New Jersey:

Shepard, L. A. (2000). The role of classroom in teaching and learning (CSE Tech. Report 517*).* Los Angeles, CA: University of California.

Singer, J. D. & Willet, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event standards-based instruction. (4th ed.).* Boston, MA: Allyn & Bacon.

Suah, S. L., & Ong, S. L. (2012). *Investigating assessment practices of in-service student learning: Doing it right-using it well.* Assessment Training Institute: London.

Suzieleez, S., Venville, G. & Chapman, A. (2009) 'Classroom Assessment: Juxtaposing teachers' beliefs with classroom practices'. *Australian Association for Research in Education: International Education Research Conference*, 29 November to 3 December. Available at: http://www.academia.edu/421177 /Classroom_Assessment_Juxtaposing_Teachers_Beliefs_with_Classroom_Practices (Accessed: 8 November 2013).

Tabachnick, B.G. & Fidell, L. S. (2001). Using Multivariate Analysis. 4th Edition. Boston: Allyn and Bacon.

Walshaw (2012). Using computer based assessments in a large statistics service teachers. *International online journal of educational science*, 4(1), 91-106.

Zhang, Z. & Judith, A. B. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education,* 16(4), 323–342.