EXAMINING THE RELIABILITY AND VALIDITY OF RESEARCH INSTRUMENTS USING PARTIAL LEAST SQUARES STRUCTURAL EQUATION MODELING (PLS-SEM)

Sane Hwui Chan,
Yoon Fah Lay

Introduction

Today, global education systems face enormous challenges to nurture young learners who will thrive in the rapid-paced and knowledge-intensive digital world. Globalisation, technology advancements, and shifting economic patterns have become the driving forces which caused disruptive paradigm shifts. Higher expectations from global society and greater needs among the students pressurized teacher to be more versatile in the classroom. The exponential change has affected science education in Malaysia where reform evident has been quite obvious. In addition, these shifts have shown important implications on the teachers' profession development as teacher's role is not only acted as the mediator in refining Malaysian education system but also the most significant agent in this reform.

Problem of Research

In Malaysia, transformation occurred in stages since 2011 for primary school level, whereas secondary schools were initiated in 2017 when Ministry of Education (MOE) substituted the previous Secondary School Integrated Curriculum (KBSM) with the Secondary School Standard Curriculum (KSSM). The revised of existing science curriculum was prompted by the deteriorating performance of Malaysian students in TIMSS and PISA since the past eighteen years. In relation to this, Malaysian teachers must continuously adapt and respond to current changes, so that they can produce learners who are proactive and capable to strive in a high competent society. TIMSS report in 2015 disclosed that teachers' readiness is one of the prominent factors contribute to the improvement of students' science performance (Kementerian Pendidikan Malaysia, 2016). By understanding human behavioural intention, guided primarily by the theory of reasoned action (Fishbein & Ajzen, 1975) and the theory of planned behaviour (Ajzen, 1991), teachers' behaviour in the classroom can be predicted. Nevertheless, many of the previous studies were measured and analysed by using first generation multivariate analysis

Abstract. *Understanding teacher's behavioural intention in teaching science becomes a foundation to improve teacher education programme. Prior to the evaluation of the interrelation between teacher self-efficacy beliefs, teaching motivation, attitudes towards teaching science, and behavioural intention in teaching science, this research examines the reliability and validity of instruments used to measure the constructs. 'Science Teaching Efficacy Belief Instrument-Form B' (STEBI-B) was used to measure pre-service science teachers' self-efficacy beliefs. 'Work Tasks Motivation Scale for Teachers' (WTMST) was used to measure teaching motivation. 'Dimensions of Attitude towards Science' (DAS) was used to measure attitudes towards teaching science whereas the 'Behavioural Intention Scale' was used to measure behavioural intention in teaching science among pre-service science teachers. Partial Least Square-Structural Equation Modelling approach was used to evaluate the reliability and validity of the instruments. Research findings concluded all instruments are valid and reliable to be used in future research.*

Key words: *self-efficacy beliefs, teaching motivation, attitudes towards teaching science, behavioural intention, partial least square-structural equation modeling.*

Sane Hwui Chan
Kian Kok Middle School, Malaysia
Yoon Fah Lay
University Malaysia Sabah, Malaysia

methods which often disregard the measurement error and are less accurate. Although these instruments have been used for more than few decades, there have been relatively few attempts to validate the research instruments. In fact, literature which involved in validating research instruments using second generation data analysis is still scarce. Thus, prior to evaluate the interrelation between constructs that predict teachers' behavioural intention in teaching science, it is crucial to examine the reliability and validity of the research instruments.

Research Focus

In consonance with Tschannen-Moran et al. (1998), teachers with self-efficacy beliefs possessed higher motivation, were in control and often confident that they can influence their students' achievement. Several authors have stipulated the effect of teacher self-efficacy beliefs (TSEB) towards students' engagement and academic performance in the classroom (Çakiroglu, Çakiroglu, & Boone, 2005; Caprara, Barbaranelli, Steca, & Malone, 2006; Capri, Ozkendir, Ozkurt, & Karakus, 2012; Goddard, Hoy, & Hoy, 2000; Kurt, Güngör, & Ekici, 2014; Mohamadi & Asadzadeh, 2012; Mojavezi & Tamiz, 2012; Senler & Sungur, 2010; Tschannen-Moran et al., 1998). In the theory of self-efficacy, Bandura postulated that beliefs can strongly influence how much effort a person willing to put forth, how determined is their endurance in the face of dispute, how resilient they are in coping with occurrence which indirectly affected their behavioural intention to teach (Bandura, 1994). Thus, teachers who possessed high self-efficacy beliefs are often more mind receptive to new teaching strategies and willing to implement different pedagogical approach to better fulfill the students' requirements (Berman, McLaughlin, Bass, Pauly, & Zellman, 1977; Tschannen-Moran et al., 1998).

Beyond that, teaching motivation (TM) emerges to be another important affective domain which influences one's working performance (Gagné & Deci, 2005). It is grounded in self-determination theory (Ryan & Deci, 2000). Teaching motivation has impact on students' academic performances as it is strongly associated with the effective teaching styles adapted by teachers in the classroom (Hein, 2012; Oredein & Awodun, 2013). Teacher who shows negative affective reactions such as low teaching motivation may result in the avoidance to teach science and therefore manifest poor commitment in teaching (Cavallo, Miller, & Saunders, 2002). Other than this, attitudes towards teaching science (ATTS) are vital in students' learning outcome preliminary in secondary level of education where students are being assessed in TIMSS and PISA. Based on OECD (2014), teachers who practice cognitive-activation strategies into their pedagogical approach can enhance students' perseverance and acceptance towards varied learning styles. Furthermore, beliefs and attitudes are two essential constructs which were found to be predictive towards the behavioural intention in teaching science (BIITS) (Denessen, Vos, Hasselman, & Louws, 2015; Lumpe, Haney, & Czerniak, 2000; Osborne, Simon, & Collins, 2003). This research is vital due to the scarcity of research emphasized on the interrelation between teacher self-efficacy beliefs, teaching motivation, attitudes towards teaching science, and behavioural intention in teaching science amongst pre-service science teachers in Malaysia. Despite of the promising results suggesting the relation between the constructs, there is no distinct evidence of the construct validation for the research instrument by using Partial Least Squares Structural Equation Modelling (PLS-SEM) approach. Considering the need for further research investigating the interrelation between the constructs, the objective of this research is to examine the reliability and validity of the instruments used in measuring teacher self-efficacy beliefs, teaching motivation, attitudes towards teaching science, and behavioural intention in teaching science.

Methodology of Research

General Background

This research has adopted quantitative research design with cross-sectional survey approach and using second–generation data analysis that is the PLS-SEM approach. The science teaching efficacy beliefs in this research were measured by the 'Science Teaching Efficacy Belief Instrument – Form B' (STEBI-B). It measures 'Personal Science Teaching Efficacy' (PSTE) and 'Science Teaching Outcome Expectancy' (STOE) amongst pre-service science teachers (Enochs & Riggs, 1990). 'Work Tasks Motivation Scale for Teachers' (WTMST) is another instrument adapted into this research to measure teaching motivation in terms of 'Intrinsic Motivation', 'Identified Regulation', 'Introjected Regulation', External Regulation', and 'Amotivation' among pre-service science teachers (Fernet, Senécal, Guay, Marsh, & Dowson, 2008). 'Dimensions of Attitude towards Science' (DAS) instrument was designed to measure pre-service science teachers' attitudes towards teaching science in terms of 'Relevance of Teaching Science' (R), 'Difficulty if

Teaching Science' (D), 'Gender Stereotypical Beliefs regarding Teaching Science' (G), 'Enjoyment in Teaching Science' (E), 'Anxiety in Teaching Science' (A), 'Self-Efficacy' (S), and 'Perceived Dependency on Context Factors' (C) (Van Aalderen-Smeets & Walma van der Molen, 2013). The behavioural intention in teaching science was measured by using the 'Behavioural Intention Scale' (Van Aalderen-Smeets & Walma van der Molen, 2013).

Sample

Sample respondents in this research were randomly selected based on cluster sampling technique from a population of 20 government-funded universities in Malaysia. The Malaysian Ministry of Higher Education (MOHE) has categorized government-funded universities into three clusters. In this research, three universities which represent each cluster of university were randomly selected namely University of Science, Malaysia (USM), Sultan Idris Education University (UPSI), and University Malaysia Sabah. The target respondents were chosen among the pre-service science teachers who are currently studying sixth semester in respective universities. Actual population size for the distribution of sampled respondents based on corresponding universities is shown in Table 1. A total of 60 sample respondents were selected randomly for pilot research with 20 respondents representing each of the three universities.

Table 1. **Distribution of sampled respondents who are currently studying in sixth semester according to universities.**

Universities	n
University of Science, Malaysia (USM)	20
Sultan Idris Education University (UPSI)	20
University Malaysia Sabah (UMS)	20
Total	60

Instrument and Procedures

The research instrument includes Science Teaching Efficacy Belief Instrument-Form B (STEBI-B), Work Tasks Motivation Scale for Teaching (WTMST), Dimensions of Attitudes towards Science (DAS) Instrument, and Behavioural Intention Scale.

Science Teaching Efficacy Belief Instrument-Form B (STEBI-B)

The STEBI-B instrument used in this research was adapted from the Science Teaching Efficacy Belief Instrument – Form A (STEBI-A), which was initially formulated for in-service science teachers and was theoretically based on Bandura's Social Learning Theory (Bandura, 1971). It consists of 13 items which measure the 'Personal Science Teaching Efficacy Belief' (PSTE) dimension and 10 items were developed to measure 'Science Teaching Outcome Expectancy' (STOE) dimension. According to Enochs & Riggs (1990), the reliability analysis reported to show high loaded factor analysis with alpha coefficient of .90 and .76 for PSTE and STOE, respectively.

Work Tasks Motivation Scale for Teachers (WTMST)

WTMST was designed grounded on Self-Determination Theory (Ryan & Deci, 2000) to evaluate five motivational regulations toward specific work tasks, which includes intrinsic motivation, identified regulation, introjected regulation, external regulation, and amotivation. According to Fernet et al. (2008), the WTMST was designed to be versatile in which researcher can adapt the instrument for specific work tasks that are relevant to the research objective. Hence, this study will adapt the items used to measure teaching task from among the recommended six work tasks. A total of 15 items assessing teaching motivation work tasks were selected out of 90 items. Each of the five motivational regulations was reflected by three items. Research findings by Fernet et al. (2008) show the internal consistency for the instrument has met criterion of .70 and confirmatory factor analysis (CFA) results have shown a good fit with salient loadings higher than .50.

Dimensions of Attitudes towards Science (DAS) Instrument

The DAS Instrument was developed to measure teachers' attitudes towards teaching science (Van Aalderen-Smeets & Walma van der Molen, 2013). It was designed based on a new theoretical framework that describes different dimensions of a primary teacher's teaching attitude, i.e. his/her attitude towards teaching science. The framework is composed of three dimensions (Cognition, Affect, and Perceived Control) with seven subdimensions related to thoughts, beliefs, and feelings towards teaching science. The DAS instrument includes a total of 28 items, composed of seven subscales which are 'Relevance of Teaching Science' (R), 'Difficulty of Teaching Science' (D), 'Gender Stereotypical Beliefs regarding Teaching Science' (G), 'Enjoyment in Teaching Science' (E), 'Anxiety in Teaching Science' (A), 'Self-Efficacy' (S), and 'Perceived Dependency on Context Factors' (C). In consonance with Van Aalderen-Smeets & Walma van der Molen (2013), the CFA showed excellent result with factor loadings varying between .38 and .90 and the internal consistency proven to be distinct with Cronbach's alpha values range from .74 to .93.

Behavioural Intention Scale

Behavioural Intention Scale was developed by Van Aalderen-Smeets & Walma van der Molen (2013) to predict pre-service and in-service teachers' behaviour in teaching science. It consisted of eight items to measure four subscales which are self-efficacy (SE), context dependency, enjoyment, and anxiety. The multiple coefficients of this scale reported to be 0.46 with R square value showing 0.214.

The administration of research instruments began once all the formal permission was obtained from the related authorities. Prior to the administration of the survey questionnaire, sample respondents were given consent letters to obtain their consensus before they agree to participate in the research. This is to ensure the feedback regarding the research is thoroughly genuine. They were informed about the objective of this research and clear instructions were given on how the instruments should be responded. Respondents were required to indicate their degree of agreement and disagreement for each of the items contained in the questionnaire. In the context of this research, modification was done to alter the Likert scale into ten-point semantic differential scales range from 'Strongly Disagree' (score 1) to 'Strongly Agree' (score 10) for STEBI-B and DAS Instrument. In WTMST, the ten-point semantic differential scales range from 'Does Not Correspond at All' (score 1) to 'Correspond Completely' (score 10). For items in the Behavioural Intention Scale, respondents were asked to indicate their perceived intention to engage in science related activities on a modified ten-point semantic differential scales range from 'Seldom or Never' (score 1) to 'Daily' (score 10).

Data Analysis

Before running PLS-SEM analysis, the data collected were screened to ensure error free from missing value, suspicious response patterns, and outliers. Results from the statistical analysis were reviewed and evaluated in terms of the relation among items in the measurement model. In current research, the internal consistency reliability, convergent validity, and discriminant validity for each individual item of the instruments were being assessed. Internal consistency reliability for each subscale was determined by the composite reliability and Cronbach's alpha coefficient, meanwhile the Average Variance Extracted (AVE) was being evaluated to assess the convergent validity of the instrument. Fornell-Larker criterion, cross-loadings, and Heterotrait-Monotrait ratio (HTMT) were assessed to evaluate the discriminant validity for each item in the instrument.

Results of Research

In Appendix 1, total_B, total_C, total_D, and total_E are the composite items used to measure latent constructs for TSEB, TM, ATTS, and BIITS, respectively. In the normality tests for reflective items that measured TSEB, the skewness value is .459 which indicates positively skewed, whereas the kurtosis value is -.327 which indicates the distribution is platykurtic. As for reflective items that measured TM, the skewness value is .832; meanwhile the kurtosis value is .825. For ATTS, the skewness and kurtosis values for the reflective items are .709 and .825, respectively. Besides, for items that measured BIITS, the skewness value is -.483 indicates negatively skewed, while the kurtosis value is .725. Based on the results, it can be summarized that the data falls within the normal distribution range. For the Kolmogorov-Smirnov test shown in Table 2, the significant level is reported to be .002 ($p < .05$) for total B, .056 (p

> .05) for total C, .032 ($p < .05$) for total D, and .200 ($p > .05$) for total E. Results show that the data is normally distributed for latent constructs teaching motivation and behavioural intention in teaching science, nonetheless it is non-normal for the latent constructs teacher self-efficacy beliefs and attitudes towards teaching science.

Table 2.   Kolmogorov-Smirnov normality test.

	Kolmogorov-Smirnova		
	K-S	df	p
Total_B	.149	60	.002
Total_C	.113	60	.056
Total_D	.120	60	.032
Total_E	.080	60	.200*

*. This is a lower bound of the true significance ($p < .05$).

a. Lilliefors Significance Correction
K-S. Kolmogorov-Smirnov

Internal Consistency Reliability for STEBI-B, WTMST, DAS Instrument, and Behavioural Intention Scale

In this research, the reflective measurement model for the instruments was being evaluated to measure the internal consistency reliability. Based on Figure 1, the composite reliability for all four constructs is above the minimum threshold value (.70). Furthermore, the matrix tab in Table 3 shows the composite reliability value appeared to be .916 (ATTS), .884 (BIITS), .881 (TM), and .901 (TSEB). Thus, current research findings show the instruments have achieved great degree of internal consistency reliability.
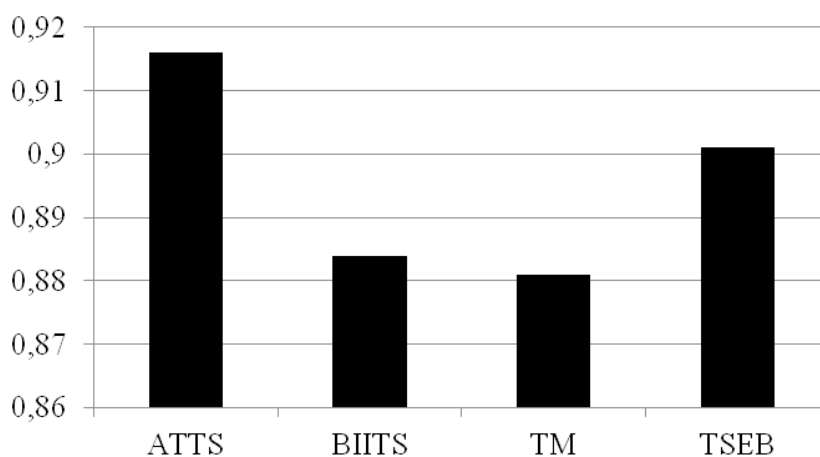

Figure 1:   Composite reliability to show the construct reliability and validity.

Table 3.   The internal consistency reliability of the instruments based on construct TSEB, TM, ATTS, and BIITS.

Matrix	Composite Reliability	Cronbach's Alpha	Average Variance Extracted
ATTS	.916	.887	.648
BIITS	.844	.786	.531
TM	.881	.821	.650
TSEB	.901	.877	.503

Cronbach's alpha (CA) was another criterion used to measure how well each individual item in a subscale correlated with the sum of the remaining items with threshold value of .70 (Hair, Hult, Ringle, & Sarstedt, 2017). All bars in the chart were above the threshold value as shown in Figure 2. Based on Table 3, the specific Cronbach's alpha reported to be .887 for ATTS, .786 for BIITS, .821 for TM and .877 for TSEB. From the results of the composite reliability and Cronbach's alpha, it shows that the research instruments are reliable survey tools which are able to measure complex affective domains for all the constructs.
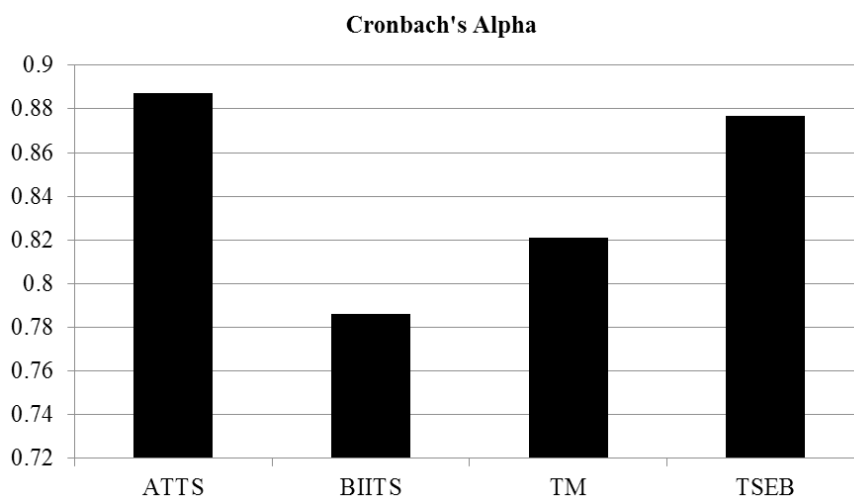

Figure 2:   Cronbach's alpha to show construct reliability and validity.

Convergent Validity of STEBI-B, WTMST, DAS Instrument, and Behavioural Intention Scale

Table 2 shows the Average Variance Extracted (AVE) value of ATTS (.648) is beyond the required lowest threshold value of .50. This is also applicable to BIITS (.531), TM (.650), and TSEB (.503). By referring to Figure 3, all the bar charts clearly indicating these four reflective constructs have achieved the minimum threshold value and therefore the instruments used in measuring the four constructs have established high levels of convergent validity.
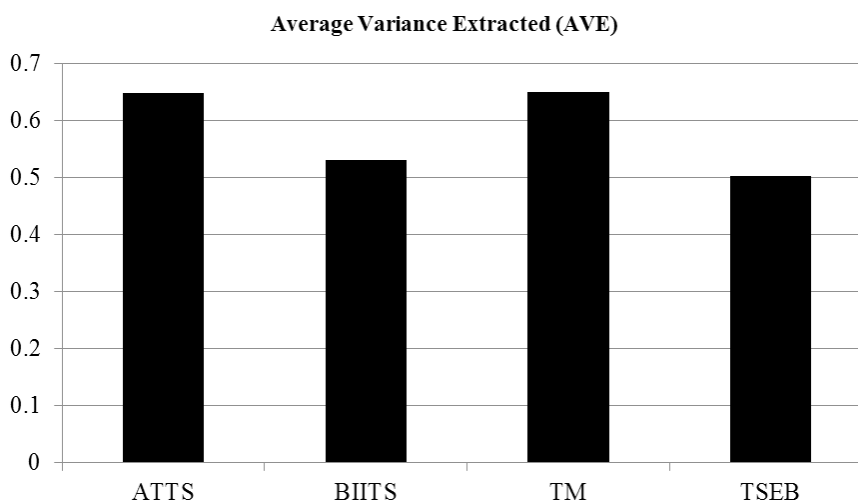

Figure 3:   Average Variance Extracted (AVE) to show construct reliability and validity.

Discriminant Validity of STEBI-B, WTMST, DAS Instrument, and Behavioural Intention Scale

Table 4 shows the results of the Fornell-Larcker criterion assessment with the reflective construct ATTS has a value of 0.805 for the square root of its AVE. This value is higher than the BIITS (.573), TM (.721), and TSEB (.586). As for the reflective construct of BIITS, it has a value of .729 for the square root of its AVE which is greater than TM (.353) and TSEB (.424). The square root of AVE for TM is .806 which is greater than TSEB (.722) whereas the reflective construct for TSEB has a value of .709 for the square root of its AVE.

Table 4. Fornell-Larker Criterion for the constructs TSEB, TM, ATTS, and BIITS.

	ATTS	BIITS	TM	TSEB
ATTS	.805			
BIITS	.573	.729		
TM	.721	.353	.806	
TSEB	.586	.424	.722	.709

Table 5 shows the cross-loadings for each indicator which reflected on four different latent constructs (i.e., ATTS, BIITS, TM, and TSEB). Items B11, B12, B14, B15, B16, B18, B22, B4, and B9 load high on its corresponding construct TSEB and much lower on other constructs ATTS, BIITS, and TM. Items C10, C15, C3, and C4 load high on its corresponding construct TM and also lower on other constructs ATTS, BIITS, and TSEB. Items D_C2, D_E1, D_E2, D_E3, and D_E4 also appeared to load high on its corresponding construct ATTS but much lower on other constructs BIITS, TM, and TSEB. Furthermore, items E_PS4, E_PS5, E_PS6, E_PS7 and E_PS8 also load higher on its corresponding construct BIITS and lower on other constructs ATTS, TM, and TSEB. From the research findings, it shows that all loadings exceeded the cross-loadings in which this indicates the discriminant validity has been established for all four constructs. Meanwhile, Figure 4 illustrates all the different outer loadings of the indicator corresponding to respective latent constructs in the measurement model. Most of the indicators retained in the measurement model have outer loadings values exceeded the threshold value of .708. Other items including B11, B16, B18, B4, D_R3, E_PS6, and E_PS8 were not removed from the instruments as the AVE value is still well above the minimum requirements with the inclusion of the items.

Table 5. Cross loadings for the constructs TSEB, TM, ATTS, and BIITS.

Items	ATTS	BIITS	TM	TSEB
B11	0.426	0.290	0.467	0.700
B12	0.386	0.306	0.457	0.743
B14	0.306	0.189	0.499	0.741
B15	0.398	0.196	0.456	0.741
B16	0.264	0.202	0.466	0.622
B18	0.354	0.373	0.510	0.676
B22	0.423	0.504	0.554	0.713
B4	0.360	0.076	0.528	0.647
B9	0.492	0.364	0.632	0.782
C10	0.489	0.332	0.773	0.640
C15	0.631	0.285	0.869	0.629
C3	0.435	0.249	0.780	0.564
C4	0.556	0.271	0.799	0.494
D_C2	0.678	0.324	0.666	0.482
D_E1	0.847	0.516	0.567	0.476

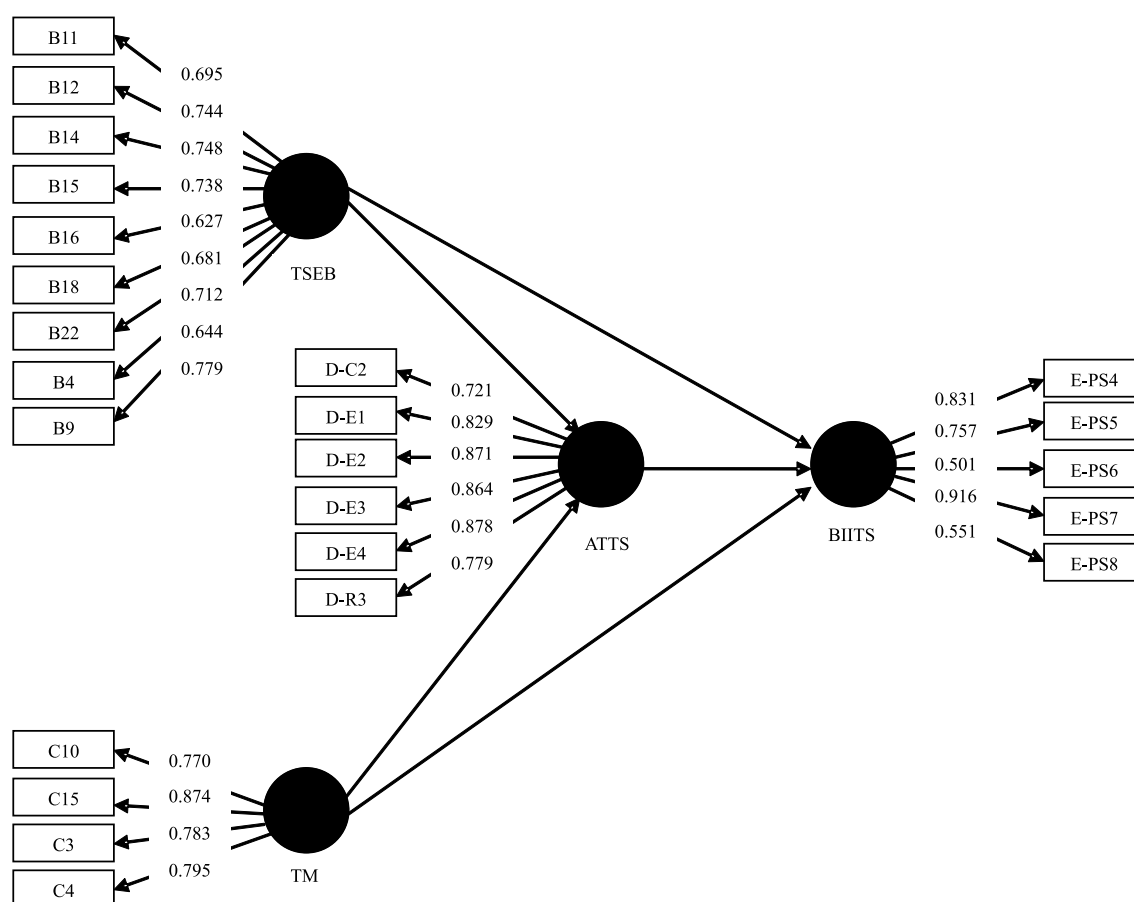| Items	ATTS	BIITS	TM	TSEB
D_E2	0.910	0.562	0.485	0.384
D_E3	0.881	0.527	0.557	0.500
D_E4	0.906	0.522	0.535	0.474
E_PS4	0.523	0.836	0.286	0.424
E_PS5	0.293	0.763	0.137	0.141
E_PS6	0.227	0.496	0.197	0.140
E_PS7	0.622	0.916	0.411	0.470
E_PS8	0.186	0.534	0.093	0.097

Figure 4:   Outer loadings and R² values for the measurement model.

The Heterotrait-Monotrait Ratio (HTMT) was shown in Table 5 and bar charts in Figure 5. All the constructs fall under the maximum threshold value of .85 in which all the constructs in the proposed path model are conceptually more distinct.

Table 5.   Heterotrait-Monotrait Ratio (HTMT).

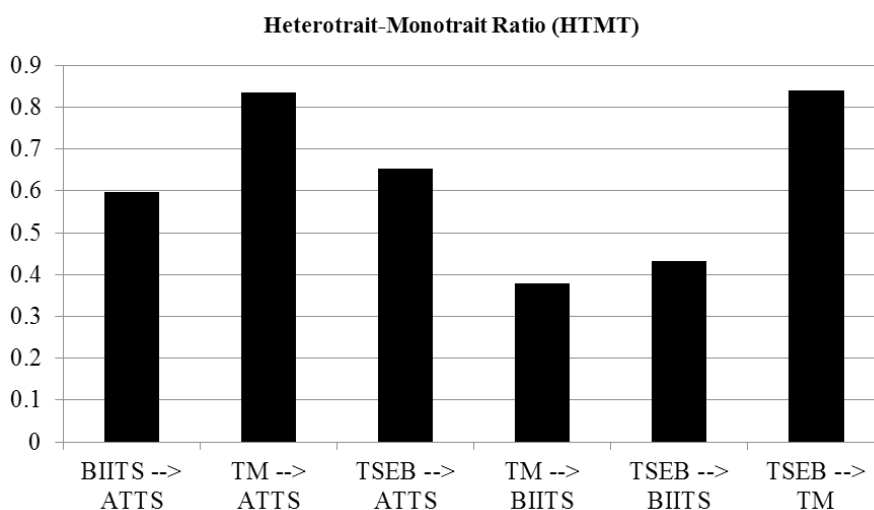|	ATTS	BIITS	TM	TSEB
ATTS				
BIITS	.597			
TM	.836	.380		
TSEB	.653	.433	.841	

Figure 5:   Heterotrait-Monotrait Ratio (HTMT) to show the discriminant validity.

Furthermore, the bar charts shown in Figure 5 also clearly indicate the discriminant validity has been established. From the bar charts, it illustrates the HTMT for BIITS à ATTS is .597, TM à ATTS is .836, TSEB à ATTS is .653, TM à BIITS is .380, TSEB à BIITS is .433, and TSEB àTM is .841.

Discussion

The aim of this research was to examine the reliability and validity of research instruments by using the PLS-SEM approach. It is crucial to ensure the instruments have achieved reliable and valid measures of constructs prior to drawing conclusion regarding the interrelation of the constructs. Although many questionnaires have been developed in previous literature, there is still scarce of research which focused on validating instruments in multidimensional. In previous research, the validation for research instruments (i.e. STEBI-B, WTMST, DAS Instrument, and Behavioural Intention Scale) were mainly emphasized on the Cronbach's alpha and CFA values by using first generation statistical analysis approach. The instruments were not validated in multidimensional, and hence measurement error has been neglected. Thus, this research will look into different aspects of validation, such as the internal consistency reliability, convergent validity, and discriminant validity for each individual item of the instruments. With the second generation statistical analysis, PLS-SEM approach uses different indicators to measure the reliability and validity of the instruments. The revalidation of instruments by using more sophisticated approach is capable to strengthen the precision of the instruments in measuring specific constructs. The improved accuracy is based on the assumption that using several indicators to measure the instrument is able to represent the various aspects of validation.

According to Enochs & Riggs (1990), the reliability analysis for STEBI-B which is used to measure TSEB reported to show Cronbach's alpha value of .90 and .76 for PSTE and STOE, respectively. As in the findings of this research, the composite reliability and Cronbach's alpha value appeared to be .901 and .877. The AVE value for the TSEB is .503 which is well above the minimum level of .50 and this showed that STEBI-B has established distinct levels of convergent validity. For the discriminant validity, Fornell-Larker Criterion for TSEB is .709 for the square root of its AVE. The initial 23 items of STEBI-B were subjected to measure its discriminant validity. Nine items (i.e., B4, B9, B11, B12, B14, B15, B16, and B18) were remained in STEBI-B due to high factor loadings. All items in STEBI-B also fall under the maximum threshold value of .85 for the HTMT which clearly showed the instruments are conceptually more distinct. Research findings by Fernet et. al. (2008) show the results for internal consistency for WTMST of .70 and CFA results have shown salient loadings higher than .50. In comparison with the findings of this research, the composite reliability value appeared to be .881 and Cronbach's alpha is .821 which indicates distinct internal consistency reliability. The AVE value for TM is .650 which is also above the threshold value of .50. Fornel-Larker Criterion for TM is .806 for the square root of its AVE. Among the 15 items of WTMST, four items (i.e., C3, C4, C10, and C15) were remained in WTMST due to high factor loadings and HTMT value lower than 0.85.

Van Aalderen-Smeets & Walma van der Molen (2013) developed the DAS instruments and they reported the CFA showed excellent values which varied from .38 to .90 and Cronbach's alpha values range from .74 to .93. As for this research, the findings are compatible with the one reported in previous study. Both composite reliability and Cronbach's alpha values are reported to be high which are .916 and .887, respectively. Furthermore, DAS instrument has also established high level of convergent validity with the AVE value of .648. Fornel-Larker criterion and HTMT values proven all the items in DAS instrument were highly discriminate. Amongst the 28 items, six items (i.e., D_C2, D_E1, D_E2, D_E3, D_E4, and D_R3) were remained in DAS instrument.

The Behavioral Intention Scale which developed by Van Aalderen-Smeets & Walma van der Molen (2013) only reported the multiple coefficients of this scale which is .46 with R square value showing .214. The internal consistency reliability based on PSL-SEM analysis in this research reported the composite reliability is .844, whereas the Cronbach's alpha is .887. Moreover, the AVE value for this instrument is .648 which is above the minimum requirement of .50. Fornel-Larker Criterion showed .729, whereas the HTMT for Behavioural Intention Scale also indicated high discriminant validity. Five items (i.e., E_PS4, E_PS5, E_PS6, E_PS7, and E_PS8) were remained from the initial eight items in the scale to high factor loadings.

Conclusions

The PLS-SEM statistical analysis concluded that STEBI-B, WTMST, DAS Instrument, and Behavioural Intention Scale are reliable and valid tools in measuring teacher self-efficacy beliefs, teaching motivation, attitudes towards teaching science, and behavioural intention in teaching science, respectively amongst pre-service science teachers in Malaysia. The instruments are reliable and valid to be used in future research.

References

Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behavior and Human Decision Processes*, *50*, 179–211. http://doi.org/10.1016/0749-5978(91)90020-T.

Bandura, A. (1971). Social learning theory. *Social Learning Theory*. New York: General Learning Corporation. http://doi.org/10.1111/j.1460-2466.1978.tb01621.x.

Bandura, A. (1994). *Self-Efficacy*. (V. S. Ramachaudran, Ed.) (Vol. 4). New York: Academic Press. Retrieved 3/08/2016, from http://onlinelibrary.wiley.com/doi/10.1002/9780470479216.corpsy0836/full.

Berman, P., McLaughlin, M. W., Bass, G., Pauly, E., & Zellman, G. (1977). Federal programs supporting educational change, Vol VII: Factors affecting implementation and continuation. *The Rand Corporation*, *7*, 1-238. http://doi.org/10.2172/1025774.

Çakiroglu, J., Çakiroglu, E., & Boone, W. J. (2005). Pre-service teacher self-efficacy beliefs regarding science teaching: A comparison of pre-service teachers in Turkey and the USA. *Science Educator*, *14* (1), 31–40.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44* (6), 473–490. http://doi.org/10.1016/j.jsp.2006.09.001.

Capri, B., Ozkendir, O. M., Ozkurt, B., & Karakus, F. (2012). Investigation of pre-service teachers' attitudes towards physics lesson and self-efficacy beliefs according to their gender, departments and perceived success in physics. *Procedia - Social and Behavioural Sciences*, *47*, 1034–1039. http://doi.org/10.1016/j.sbspro.2012.06.775.

Cavallo, A. M. L., Miller, R. B., & Saunders, G. (2002). Motivation and affect toward learning science among preservice elementary school teachers: Implications for classroom teaching. *Journal of Elementary Science Education*, *14* (2), 25–38.

Denessen, E., Vos, N., Hasselman, F., & Louws, M. (2015). The relationship between primary school teacher and student attitudes towards science and technology. *Education Research International*, *2015*, 1–7. http://doi.org/10.1155/2015/534690.

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *Annual Meeting of the National Association for Research in Science Teaching*, 1–30. http://doi.org/10.1300/J108v05n03.

Fernet, C., Senécal, C., Guay, F., Marsh, H., & Dowson, M. (2008). The work tasks motivation scale for teachers (WTMST). *Journal of Career Assessment*, *16* (2), 256–279. http://doi.org/10.1177/1069072707305764.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*. United States of America: Addison-Wesley Publishing Company. Retrieved 23/12/2016 from http://people.umass.edu/aizen/f&a1975.html.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*, 331–362. http://doi.org/10.1002/job.322.

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, *37* (2), 479–507.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modelling (PLS-SEM)* (2nd ed.). United States of America: SAGE Publication.

Hein, V. (2012). The effect of teacher behaviour on students' motivation and learning outcomes: A review, *18*, 9–19.

Kementerian Pendidikan Malaysia. (2016). *Laporan TIMSS 2015 - Trends in International Mathematics and Science Study*. Malaysia: Perpustakaan Negara Malaysia. Retrieved 23/12/2017, from https://www.acer.edu.au/timss.

Kurt, H., Güngör, F., & Ekici, G. (2014). The relationship among teacher efficacy, efficacy regarding teaching, and responsibility for student achievement. *Procedia - Social and Behavioral Sciences*, *116*, 802–807. http://doi.org/10.1016/j.sbspro.2014.01.301.

Lumpe, A. T., Haney, J. J., & Czerniak, C. M. (2000). Assessing teachers' beliefs about their science teaching context. *Journal of Research in Science Teaching*, *37* (3), 275–292. http://doi.org/10.1002/(SICI)1098-2736(200003)37:3<275::AID-TEA4>3.0.CO;2-2.

Mohamadi, F. S., & Asadzadeh, H. (2012). Testing the mediating role of teachers' self-efficacy beliefs in the relationship between sources of efficacy information and students' achievement. *Asia Pacific Education Review*, *13* (3), 427–433. http://doi.org/10.1007/s12564-011-9203-8.

Mojavezi, A., & Tamiz, M. P. (2012). The impact of teacher self-efficacy on the students' motivation and achievement. *Theory and Practice in Language Studies*, *2* (3), 483–491. http://doi.org/10.4304/tpls.2.3.483-491.

OECD. (2014). PISA 2012 results in focus. *Programme for International Student Assessment*, 1–44. http://doi.org/10.1787/9789264208070-en.

Oredein, A., & Awodun, A. (2013). Impact of teachers' motivational indices on science students' academic performance in Nigerian senior secondary schools. *International Education Studies*, *6* (2), 49–54. http://doi.org/10.5539/ies.v6n2p49.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25* (9), 1049–1079. http://doi.org/10.1080/0950069032000032199.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*, *55* (1), 68–78. http://doi.org/10.1037/0003-066X.55.1.68.

Senler, B., & Sungur, S. (2010). Pre-service science teacher's teaching self-efficacy: A case from Turkey. *Procedia-Social and Behavioral Sciences*, *9*, 771–775. http://doi.org/10.1016/j.sbspro.2010.12.232.

Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248. http://doi.org/10.3102/00346543068002202.

Van Aalderen-Smeets, S., & Walma van der Molen, J. (2013). Measuring primary teachers' attitudes toward teaching science: Development of the dimensions of attitude toward science (DAS) instrument. *International Journal of Science Education*, *35*, 577–600. http://doi.org/10.1080/09500693.2012.755576.

Appendix 1. **Skewness and Kurtosis Values for Composite items (Total_B, Total_C, Total_D, and Total_E) which measured TSEB, TM, ATTS, and BIITS constructs.**

Descriptives			Statistic	Std. Error
Total_B	Mean		152.32	2.029
	95% Confidence Interval for Mean	Lower Bound	148.26	
		Upper Bound	156.38	
	5% Trimmed Mean		151.83	
	Median		147.00	
	Variance		247.034	
	Std. Deviation		15.717	
	Minimum		117	
	Maximum		190	
	Range		73	
	Interquartile Range		25	
	Skewness		.459	.309
	Kurtosis		-.327	.608
Total_C	Mean		94.12	2.042
	95% Confidence Interval for Mean	Lower Bound	90.03	
		Upper Bound	98.20	
	5% Trimmed Mean		93.09	
	Median		94.50	
	Variance		250.206	
	Std. Deviation		15.818	
	Minimum		70	
	Maximum		141	
	Range		71	
	Interquartile Range		22	
	Skewness		.832	.309
	Kurtosis		.825	.608
Total_D	Mean		176.45	3.221
	95% Confidence Interval for Mean	Lower Bound	170.01	
		Upper Bound	182.89	
	5% Trimmed Mean		175.22	
	Median		170.00	
	Variance		622.421	
	Std. Deviation		24.948	
	Minimum		138	
	Maximum		233	
	Range		95	
	Interquartile Range		39	
	Skewness		.709	.309
	Kurtosis		-.235	.608

Descriptives

				Statistic	Std. Error
Total_E	Mean			52.12	1.190
	95% Confidence Interval for Mean	Lower Bound		49.74	
		Upper Bound		54.50	
	5% Trimmed Mean			52.33	
	Median			52.50	
	Variance			84.952	
	Std. Deviation			9.217	
	Minimum			23	
	Maximum			72	
	Range			49	
	Interquartile Range			12	
	Skewness			-.483	.309
	Kurtosis			.725	.608

Sane Hwui Chan	Master in Science Education, Teacher, Kian Kok Middle School, Kota Kinabalu, Sabah, Malaysia. E-mail: shanee0212@gmail.com
Yoon Fah Lay	PhD, Associate Professor of Science Education at the Faculty of Psychology and Education, University Malaysia Sabah, Malaysia. E-mail: layyoonfah@yahoo.com.my