



# European Journal of Educational Research

Volume 11, Issue 1, 179 - 191.

ISSN: 2165-8714

<https://www.eu-jer.com/>

## Development and Validation of Teachers' Assessment Belief of Mathematical Thinking Instrument

Saoud Alhunaini\* 

A'Sharqiyah University, OMAN

Kamisah Osman 

National University of Malaysia,  
MALAYSIA

Naser Abdurab 

Taiz University, YEMEN

*Received: July 4, 2021 • Revised: September 8, 2021 • Accepted: November 6, 2021*

**Abstract:** This study was conducted to corroborate in understanding the teachers' beliefs about assessment practices. The prior studies related to teachers' assessment beliefs in mathematics have been done to assess teachers' beliefs in the general context of mathematics teaching. This study developed an instrument to assess teachers' assessment beliefs of mathematical thinking. The research aimed to develop and validate a scale of assessment beliefs of mathematical thinking by using the confirmatory factor analysis. The first draft of the scale contained 25 items. The sample of the study consisted of 537 mathematics teachers from public schools in Oman. The instrument was a questionnaire with a 5-point Likert scale. The scale was validated by asking a number of experts in mathematics educational measurement and evaluation. Exploratory and confirmatory factor analysis was applied to test the model of assessment beliefs of mathematical thinking scale using AMOS 25.0. All constructs had acceptable reliability. The model had a good model fit for the assessment beliefs of mathematical thinking scale which obtainable from the fit indices tests. The findings revealed that all fit criteria indices were realized. The results also showed acceptable validity and construct reliability for the scale.

**Keywords:** *Assessment beliefs, assessment forms, assessment purpose, mathematics beliefs, mathematical thinking.*

**To cite this article:** Alhunaini, S., Osman, K., & Abdurab, N. (2022). Development and validation of teachers' assessment belief of mathematical thinking instrument. *European Journal of Educational Research*, 11(1), 179-191. <https://doi.org/10.12973/eu-jer.11.1.179>

### Introduction

Educational researchers have paid more attention to teachers' assessment beliefs and practices because they have a significant role in teachers' efforts for improving learning quality (Widiastuti et al., 2020). This motivated educator and researchers to look for ways that help to improve the classroom assessment practices and discover what factors that may increase or decrease the quality of these assessment practices like the teachers' assessment beliefs. Teachers' beliefs related to assessment are a critical portion of the teachers' system beliefs (Unal & Unal, 2019). Several researchers have examined the assessment beliefs of mathematics teachers. The findings of researches have revealed that there is a mutual effect between teachers' beliefs and their practices, especially in the classroom environment. Besides, nowadays, mathematical curricula focus on assessing mathematical thinking skills, like problem solving, reasoning, communication, mathematical communication, and representation skills (Bal & Artut, 2019). Thus, many mathematics education researchers believed that there is an important effect of teachers' assessment beliefs on their mathematics teaching and learning (Genc, 2005; Smith, 2010).

### Literature review

Assessment is used to determine and assess the students' understanding of the concepts, terms, and learning progress. It also helps to measure the level of students' acquiring of learning skills (Dandis, 2013; Siemon et al., 2017). Assessing students' understanding introduces a better view of what the students know (Unal & Unal, 2019). Assessment results are not only important for students' learning but also help teachers and educational organizations to discover the strengths and weaknesses of their educational efforts (Abdullah et al., 2019). Evaluation is intended to improve reporting of changes in teaching practices and students' learning process. It is a method of determining at which extension the

\* Corresponding author:

Saoud Alhunaini, College of Arts and Humanities, A'Sharqiyah University, Oman. ✉ [saoud.alhunaini@moe.om](mailto:saoud.alhunaini@moe.om)

© 2022 The Author(s). Open Access - This article is under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).



teaching has affected students' learning. Student accountability assessment aims to assess student performance. School accountability assessment uses student assessments and tools results to measure the quality of schools and teachers. The school accountability assessment provides information about student performance (Unal & Unal, 2019). It can be defined the students' learning and thinking strengths and weaknesses through analyzing the results of assessment tools (Ali, 2015; Haynes et al., 2016).

Researches highlighted the connection between assessment practices of teachers and their assessment beliefs. Consequently, the efforts of educational institutions promote a common understanding of the assessment criteria. However, those efforts are useless if teachers' beliefs are not taken into account because different teachers' perceptions lead to different assessment practices (Genc, 2005; Muñoz et al., 2012). Moreover, teachers' work tasks and other responsibilities in the schools could lead to a contradiction between teachers' assessment beliefs and practices. Besides, enough time is not always available for teachers to address all assessment tools and methods that are consistent with their beliefs about educational assessment (Gebril & Brown, 2014; Widiastuti et al., 2020).

Further, teachers' beliefs about assessment increase positively with the growth of their experience of teaching (Unal & Unal, 2019). Those beliefs play an important role that could play positively or negatively in the students' learning. Teachers' assessment beliefs may guide teachers toward specific methods of teaching and evaluation (Dandis, 2013; Muñoz et al., 2012; Widiastuti et al., 2020). Büyükkarcı (2014) designed an instrument to assess teachers' beliefs about formative assessment. The finding indicated that teachers held positive beliefs about formative assessment. They believed that giving students oral or written feedback without giving grades is more useful for improving the students' learning. While, Muñoz et al.'s (2012) instrument classified the teachers' assessment beliefs into teachers' beliefs responsibility of schools, development of learning and teaching, and documentation of students' learning.

#### *Assessment of mathematical thinking*

The National Council of Teaching Mathematics (NCTM) emphasized that mathematical thinking should be at the peak of teaching and assessing mathematics. Because students' mathematics learning can be fostered through improving the students' abilities in mathematical thinking (National Council of Teachers of Mathematics, 2000). Reyes-Cedeno et al. (2019) investigated the assessment practices used by teachers to assess the mathematical thinking of students in early grades. The researchers applied questionnaires and interviews to describe teachers' assessment practices of mathematical thinking. The results indicated that teachers used classroom tasks to develop students' mathematical thinking which implemented the skills: classifications, number operations, consequently the seriation and representation. They also used the thinking component which consists of metric, geometric spatial, and numeric. Teachers believed that assessment practices help them to determine the starting point for improving the students' mathematical thinking. Abdullah et al. (2019) showed that the assessment practices can be implemented to train students in high order thinking processes in mathematics spatially for problem solving.

It is important to apply a different method of assessment to measure students' mathematical thinking. Assessments should focus on the individual growth of students' skills and knowledge (Etheridge, 1989). Whereas, Büyükkarcı (2014) recommended that measuring formative assessment beliefs of teachers should cover the following:

- It is necessary to discuss the lesson objectives and learning results with the students.
- Having a clear picture of the goals for students makes them honest and responsible for evaluating themselves and their colleagues.
- Provide students with feedback that identifies the strengths and weaknesses of their answers in a way that is useful for developing their learning.

#### *Assessment beliefs instruments*

In mathematics education, some tools have been developed to assess the beliefs of mathematics teacher evaluation. For example, Unal and Unal (2019) categorized teachers' assessment beliefs into student accountability, irrelevant accountability, and improvement. Genc (2005) developed a scale for measuring beliefs related to assessment practices of mathematics teachers. The scale contained two main dimensions, teachers' beliefs about: the purposes/types of assessment, and learning perspectives on assessment. The first dimension covers the teachers' beliefs about diagnostic, summative, and formative assessments. The second dimension is learning perspectives on assessment which covers the traditional assessment and alternative assessment.

Brown (2011) introduced a model of students' assessment beliefs that contained the following dimensions: Assessment with effective benefit, irrelevance (assessment is irrelevant), improvements (assessment improves teaching and learning), and external (assessment of external factors). The model was established depending on some prior models related to studies of students' assessment conceptions. The model had examined the relationships among the main variables of the model. The findings revealed that relationships among them are significant.

The current research linked the teachers' assessment beliefs of mathematical thinking to Dandis' theory (Dandis, 2013). Dandis (2013) proposed a conceptual framework for assessment beliefs of teachers. The framework was built to get more

understanding of the relation between teachers' assessment beliefs and practices. The framework divided teachers' assessment beliefs into six dimensions. These dimensions are teachers' beliefs about: Assessment forms, assessment benefits, the connection between learning, teaching and assessment practices, students learning, assessment for learning, and assessment purpose.

Although it is important to measure the beliefs of teachers' evaluation of mathematical thinking, little is known about the beliefs of teachers' evaluation of mathematical thinking in the Arabic region, especially in Oman. This research attempts to confirm the teacher assessment practices of mathematical thinking model that contain four dimensions of assessment beliefs: teachers' beliefs about assessment forms, the benefits of evaluation, the purpose of assessment, and students' learning as is shown in Figure 1.

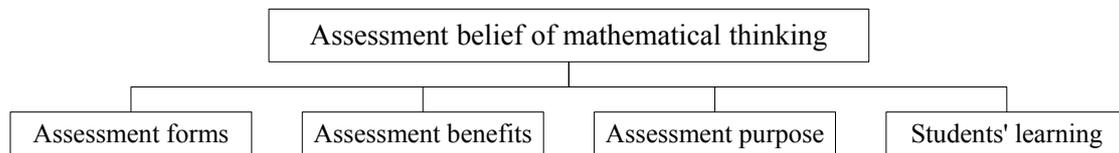


Figure 1. Assessment Beliefs of Mathematical Thinking

Assessment forms contain all kinds of assessment methods and tools such as observations, projects, direct questions, presentations, teamwork, rubric and projects. Teachers differ in their concepts of the best tools and methods for their students. Siemon et al. (2017) found that teachers used multiple assessment forms because they believed that using multiple assessment forms helps to improve the students' understanding, and reasoning of mathematics. The results of researches revealed that using different assessments methods and tools helps teachers to gather information about students' learning of mathematics such as understanding, abilities, and achievement (Ling et al., 2012; Tsamir & Tirosh, 2006).

The beliefs about the benefits of assessment cover those beliefs related to the assessment benefits for students, teachers, and educational institutes. Where the beliefs benefits for students refer to the beliefs that are directly linked to students' learning assessment (Dandis, 2013). For instance, Smith (2010) showed that teachers believed that feedback fosters the students' mathematics learning and thinking. Further, Yong and Sam (2008) resulted that there was a positive impact of introducing feedback on students' mathematical thinking. Delice and Sevimli (2010) and Nurlu (2015) indicated students' confidence can be supported by introducing suitable feedback. Teachers' beliefs about assessment benefits also referred to teachers' beliefs about the connection between results of applying the assessment tools and the policy of educational organizations. For instance, there are some teachers' beliefs that connect between the results of assessment tools and their institute (Calveric, 2010). Furthermore, the beliefs about the assessment benefits can also be connected to teachers' beliefs about how assessment helps the teachers to teach and achieve the educational objectives teaching (Youmans et al., 2018).

Teachers' beliefs about the assessment purpose contain their beliefs that relate to the objectives of teachers for assessing their students. These beliefs can illustrate the reasons that lead teachers to apply some assessment practices and other practices are excluded. In other words, teachers may depend on their purpose of assessment for their decisions for addressing the assessment methods and tools. (Büyükkarçı, 2014; Calveric, 2010; Dandis, 2013; Muñoz et al., 2012). Widiastuti et al. (2020) found that teachers believed that the purpose of using formative assessment is to improve students' learning by determining the students' weak points of learning. Reyes-Cedeno et al. (2019) indicated that teachers applied assessment as a diagnostic assessment to determine the students' experience about topics.

Teachers' beliefs related to the students' learning comprise the beliefs about the students' characteristics or abilities of learning. For example, teachers believe that it is not reasonable to implement the same assessment tools at the same level to all students because students differ in their characteristics and capabilities. In addition, there are teachers who believed that students can solve problems mechanically without connecting addressing them within the real world. Some other teachers believe that exams motivate students to learn mathematics. However, teachers are not the same in their conceptions about the students' abilities of learning (Calveric, 2010; Dandis, 2013). Youmans et al. (2018) revealed that teachers believed that students need to be supplied with assessment games for students in the early learning years.

However, most previous researches that studied mathematics teachers' assessment beliefs have focused on teachers' assessment beliefs of mathematics learning in general, and little is known about instruments for teachers' assessment beliefs of mathematical thinking. Besides, there is little known about instruments in the Arab world that have been developed for measuring teachers' assessment beliefs of mathematical thinking and tested with descriptive statistics, normality, reliability, and validity. Thus, the current research aimed to:

- Develop an instrument to be used for measuring teachers' assessment beliefs of mathematical thinking.
- Test the reliability of SEM of teachers' assessment beliefs of mathematical thinking scale.
- Test the validity of SEM of teachers' assessment beliefs of mathematical thinking scale.

- Test the G factor of SEM of teachers' assessment beliefs of mathematical thinking scale.

## Methodology

### Research Design

The current research design was conducted as a survey with the first and second-order factor structure model as was stated by Zainudin et al. (2018). The research applied a quantitative tool that is widely used for social science researches. The instrument contains a series of items that require to be responded to by candidates. It is usually called a self-reported or questionnaire because candidates generally respond to questions or statements related to beliefs, themselves, attitudes, and so on (Privitera, 2020).

### Sample of the Study

In the Omani public schools, all teachers use the same textbooks and apply the same assessment system which was designed by the Ministry of Education in Oman. The sample of this research was randomly selected from a population of mathematics teachers from Omani public schools that belong to the Ministry of Education in Oman. The researchers randomly selected six out of eleven educational governorates in the Sultanate of Oman. The researchers made sure that the selected educational governorates cover all the environments in Oman to make sure that the sample presented all teachers' demographic backgrounds in Oman. They met some mathematics education supervisors who agreed to help the researchers in applying the instrument and explained how the instrument will apply. Then schools were randomly selected by mathematics education supervisors. The sample contained 537 mathematics teachers from both males (265) and females (272). Their teaching experience is from one year to twenty-seven years. They were requested to give their opinion in the instrument of the Assessment Beliefs of Mathematical Thinking (ABMT).

### Instrument

The instrument of Assessment Beliefs of Mathematical Thinking (ABMT) was designed as a self-reported questionnaire comprise of items on a 5-point Likert scale ranging as (5=strongly agree, 4=agree, disagree, 3= neutral, 2=disagree, and 1=strongly). The items were published within three stages. The first stage was analyzing some instruments of teachers' assessment beliefs that were developed in earlier studies, both within contexts and in general educational assessment (Calveric, 2010; Gebril & Brown, 2014; Genc, 2005). The second stage was analyzing the standard of the National Council of Teachers of Mathematics (2000) for school mathematics relating to the skills of mathematical thinking. The standards were classified related to the main skills of mathematical thinking to be linked with teachers' assessment beliefs. In the third stage, the items were designed by merging the standards with the assessment beliefs items that were drawn from previous items of assessment beliefs scales. The items were written in the Arabic language to be applied with mathematics teachers who teach mathematics in the Arabic language. Then, the items were classified into four factors (Assessment Forms, Assessment Benefits, Assessment Purpose, and Students Learning) as shown in Table 1.

Table 1. The Number of Items in Each Dimension of Assessment Beliefs of Mathematical Thinking Instrument

Assessment beliefs of mathematical thinking dimensions	Number of items
Assessment Forms	5
Assessment Benefits	5
Assessment Purpose	5
Students Learning	5
Total of items	20

Content validity of the instrument items was assessed by a group of experts in the areas of mathematics education, educational evaluation, and educational measurement. Items content validity index (I-CVI) was calculated for each item by dividing the number of experts giving a rating of "relevant" by the total number of them. The experts were asked to check the clarity of language, wording, and suitability of items to be included in the instrument of ABMT for the mathematics teachers in Oman. They also denoted the importance and comprehensiveness of the items. Their feedbacks were addressed to edit the items. The experts approved that most of the items were clearly and appropriately expressed for the contestants and related to the school mathematics content from grades 5 through 12 and associated with the constructions being assessed. The results of I-CVI showed that most of the items got I-CVI > 0.79 which mean they were accepted. However, they suggested little modifications in the expressions of some items. Five items got I-CVI between 0.70 and 0.79 which means they needed to be reviewed and made the changes as suggested by the experts. All dimensions remained at the same number of items. After that, the instrument was given to a team of mathematics senior teachers and mathematics supervisors who work in the Omani public schools to check the instrument suitability of terms included in the instrument with terms used in the Omani educational system. They approved that all items were suitable and can be understood by the teachers. Then, in the second stage, the researchers randomly selected 92 mathematics teachers from public schools for applying the instrument to test the validity and reliability. Testing the instrument included taking the teachers' opinions about the clearness of items.

Table 2 displays the results of the instrument reliability of the scale ABMT. The results revealed that the values of Cronbach's alpha are (0.715-0.857) for the dimensions and 0.867 for all items. The results showed that the instrument got the accepted level of reliability as mentioned by statistical researchers (Hair et al., 2013, 2014).

*Table 2. Reliability of Assessment Beliefs of Mathematical Thinking Instrument*

<b>Assessment beliefs of mathematical thinking dimensions</b>	<b>Number of items</b>	<b>Cronbach's Alpha</b>
Assessment forms	5	0.715
Assessment benefits	5	0.78
Assessment Purpose	5	0.812
Students Learning	5	0.857
<b>Total of items</b>	<b>20</b>	<b>0.867</b>

#### *Data Analysis*

The normality for the items of the instrument was checked by calculating the Skewness and Kurtosis. The factor should be placed within the satisfactory level, which is between -7 and 7 for Kurtosis, and between -3 and 3 for Skewness (Kline, 2015). The factor analysis was applied to test the baseline constructs and issue the construct validity for the ABMT. The correlation matrix among the items of the instrument ABMT was explored at the start of any steps. The recognized value for each correlation should be ( $> 0.30$ ) to be used as factor analysis of type orthogonal methods. While, factor analysis of type the oblique methods should be addressed when it ( $\leq 0.30$ ) (Tabachnick & Fidell, 2007). Confirmatory factor analysis (CFA, AMOS 22.0) was applied for construct validity for ABMT.

#### **Results**

The current research aimed to develop an instrument for teachers' assessment beliefs of mathematical thinking. Mathematics teachers were required to show their opinions for items of ABMT, which are measured by using a five-Likert point scale ranging from 5 as "strongly agree" to 1 for "strongly disagree". Table 3 shows means, standard deviations, skewnesses and kurtosises, and reliability of all items for the ABMT instrument.

The finding revealed that teachers believe that assessment required addressing feedback to improve the confidence of students in their skills of mathematical thinking. They agree that assessment should contain the implementation of various assessment forms such as self-assessment, performance evaluation, and modern technology that can be applied to the improvement of students' mathematical thinking. The results also indicated that teachers believe that applying different tools of assessment helps to reach objectivity in assessing mathematical thinking. Further, the teachers believe that they can determine the strengths and weaknesses of students' mathematical thinking skills through the assessment. Generally, the results showed that mathematics teachers believe that assessing of students' mathematical thinking helps to monitor and develop the students' mathematical thinking because using assessment tools give students an opportunity to acquire new skills. The finding also indicated that the overall mean scores of the dimensions are between 3.80 and 4.26. Thus, it can be concluded that teachers held positive and strong beliefs about assessing students' mathematical thinking.

Table 3. Mean, Standard Deviation, Skewness and Kurtosis, and Reliability of Dimensions of ABMT

Items	M	SD	Rank	Skewness $\leq \pm 3$	Kurtosis $\leq \pm 7$	CITC $\geq 0.30$	
<b>Assessment Forms (AF)</b>							
ABMT1_F1	Performance evaluation can be activated in the development of students' mathematical thinking.	4.09	0.675	5	-0.513	0.822	0.419
ABMT2_F2	It is necessary to apply various forms of assessment to evaluate students' mathematical thinking.	4.3	0.591	2	-0.26	-0.308	0.501
ABMT3_F3	The application of self-assessment to students' mathematical thinking helps to increase learning motivation.	4.29	0.649	3	-0.445	-0.323	0.587
ABMT4_F4	Feedback enhances students' confidence in their mathematical thinking skills.	4.45	0.618	1	-0.803	0.362	0.52
ABMT5_F5	Modern technology helps to improve the ways of students' training on mathematical thinking.	4.17	0.787	4	-0.635	-0.098	0.352
	Overall Mean Score	4.26	0.664			Cronbach's $\alpha = 0.710$	
<b>Assessment Benefits (AB)</b>							
ABMT6_B1	The results of the assessment can modify the methods of training students on mathematical thinking.	4.16	0.675	2	-0.533	0.692	0.417
ABMT7_B2	The current assessment system leads the teacher to teach in a way that is inconsistent with the development of students' mathematical thinking skills.	2.6	1.057	5	0.207	-0.594	0.436
ABMT8_B3	Assessment helps to measure the quality level of work provided by students.	3.97	0.779	3	-0.607	0.624	0.564
ABMT9_B4	Assessment tools provide more accuracy in feedback sessions on mathematical thinking skills.	3.94	0.796	4	-0.465	0.09	0.579
ABMT10_B5	Variety in the use of assessment tools helps to achieve objectivity in assessing students' mathematical thinking.	4.31	0.62	1	-0.55	0.582	0.454
	Overall Mean Score	3.8	0.785			Cronbach's $\alpha = 0.713$	
<b>Assessment Purpose (AP)</b>							
ABMT11_P1	Assessment used to determine the starting point for a teacher when teaching mathematical thinking skills.	4.07	0.719	3	-0.521	0.288	0.572
ABMT12_P2	Assessment is used to produce opportunities for integrating mathematical thinking skills.	4.05	0.684	4	-0.377	0.164	0.555
ABMT13_P3	Assessment is used to enable students to employ mathematical thinking skills.	4	0.718	5	-0.553	0.666	0.584
ABMT14_P4	Assessment is used to provide teachers with strengths and weaknesses in students' mathematical thinking skills.	4.24	0.733	1	-1.008	1.696	0.567
ABMT15_P5	The results of the assessment are used to review the logical sequence for teaching objectives of mathematical thinking.	4.09	0.717	2	-0.438	0.132	0.619
	Overall Mean Score	4.09	0.714			Cronbach's $\alpha = 0.798$	
<b>Students Learning (SL)</b>							
ABMT16_SL1	Assessment tools give students an opportunity to acquire new skills.	4.12	0.727	1	-0.563	0.364	0.542
ABMT17_SL2	Assessment provides students with educational opportunities to discover the correlation of mathematics to the surrounding environment.	3.97	0.811	5	-0.49	-0.097	0.681
ABMT18_SL3	Assessment trains students to use their mathematical thinking skills in their daily lives.	4.01	0.788	4	-0.62	0.412	0.7
ABMT19_SL4	Assessment reveals the level of students' ability to acquire mathematical thinking skills.	4.09	0.71	2	-0.54	0.532	0.635
ABMT20_SL5	Assessment helps students to use mathematical thinking skills to learn new topics.	4.06	0.719	3	-0.576	0.465	0.672
	Overall Mean Score	4.05	0.751			Cronbach's $\alpha = 0.841$	

Besides, the indices/indicators of Skewness belong to the interval (-3, 3) and Kurtosis belong to the interval (-7, 7), demonstrating that all items for all dimensions are located within the acceptable threshold of Skewness and Kurtosis

given by Kline (2015) which indicated that all items of ABMT are normally distributed. Moreover, the Corrected Item-Total Correlation  $\geq 0.30$  for all items with their dimensions and positively connected to their construct. The overall reliability for all dimensions is between 0.710 and 0.841, which are located at the accepted level of conventional standard for reliability.

*Confirmatory Factor Analysis for ABMT*

As it shown in Figure 2 the ABMT was evaluated by four components: Assessment forms (five items), assessment benefits (five items), assessment purpose (five items), and students' learning (five items). As displayed in Figure 2 and Table 4, the goodness of fit indices for the first order CFA of ABMT model, revealed that its fit statistics is positioned within the required criteria as Normed Chi-Squared CMINDF = 2.843 (achieved the threshold of  $< 3$ ), CFI = 0.920 (above the threshold of  $> 0.90$ ), IFI = 0.921 (passed the threshold of  $> 0.90$ ), TLI = 0.907 (achieved the threshold of  $> 0.90$ ), GFI = 0.920 (achieved the threshold of  $> 0.90$ ), RMSEA = 0.059 (achieved the threshold of  $< 0.80$ ).

Table 4. Goodness Fit Indices for ABMT

Indices	Required Scores	Baseline Model	Re-Specified Model	Second Order CFA
1 Chi-Square	-	466.171	379.159	390.726
2 DF (Degree of Freedom)	-	164	129	131
3 P (Probability)	( $<0.05$ )	0.000	0.000	0.000
4 CMINDF (Normed Chi-Square)	( $<3$ )	2.843	2.939	2.983
5 CFI (Comparative Fit Index)	( $>0.90$ )	0.920	0.931	0.928
6 IFI (Incremental Fit Index)	( $>0.90$ )	0.921	0.931	0.916
7 TLI (Tucker Lewis Index)	( $>0.90$ )	0.907	0.918	0.916
8 GFI (Goodness of Fit Index)	( $>0.90$ )	0.920	0.926	0.925
9 SRMR (Squared Root Mean Residual)	( $<0.08$ )	0.048	0.048	0.050
10 RMSEA (Root Mean Squared Error Approximation)	( $<0.08$ )	0.059	0.060	0.061

Also, the magnitude of some item loadings is less than 0.50 as the least acceptable value. Subsequently, this baseline model needs to be improved by removing items with loading below 0.50 as less theoretically contributes in shaping and modelling its respective construct. The re-specified model must capture the previous model without item (ABMT5\_F5) from the assessment forms construct and item (ABMT7\_B2) from the assessment benefits construct.

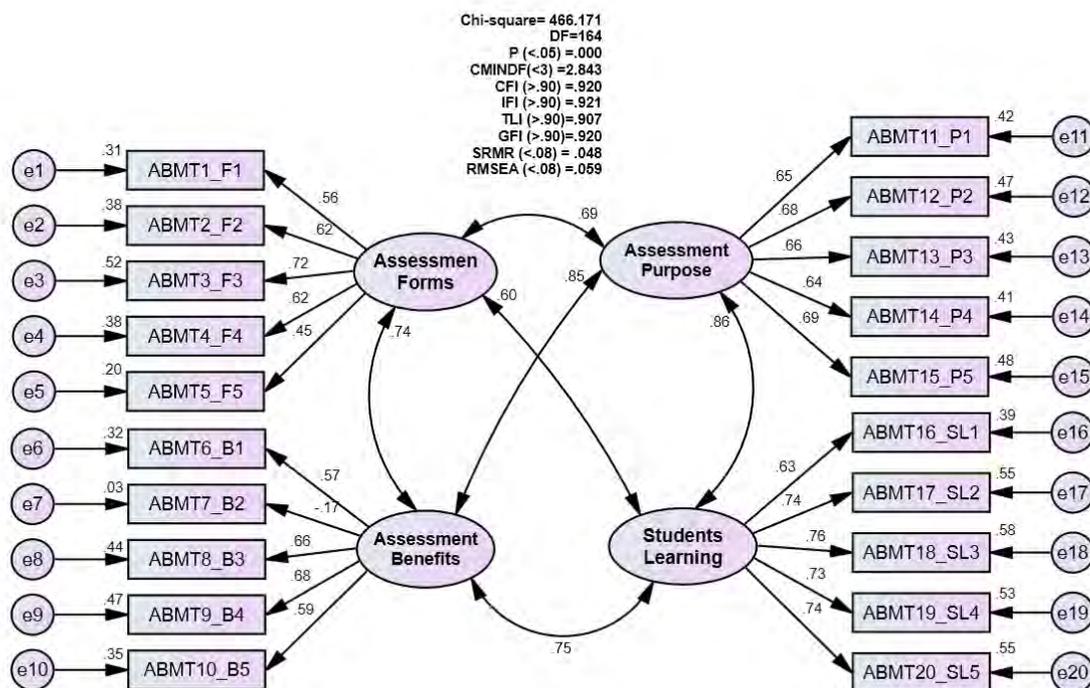


Figure 2. Confirmatory Factor Analysis for ABMT (Original Model)

As illustrated in Figure 3 and Table 4, the goodness fit indices for the first order CFA of Assessment Beliefs of Mathematical Thinking as re-specified model, exhibit that its fit statistics are located within the acceptable criteria as Normed Chi-Squared CMINDF = 2.939 (achieved the threshold of < 3), CFI = 0.931 (above the threshold of > 0.90), IFI = 0.931 (met the threshold of > 0.90), TLI = 0.918 (reached the threshold of > 0.90), GFI = 0.926 (accomplished the threshold of > 0.90), SRMR = 0.048 (met the threshold of < 0.80) and RMSEA = 0.060 (achieved the threshold of < 0.80).

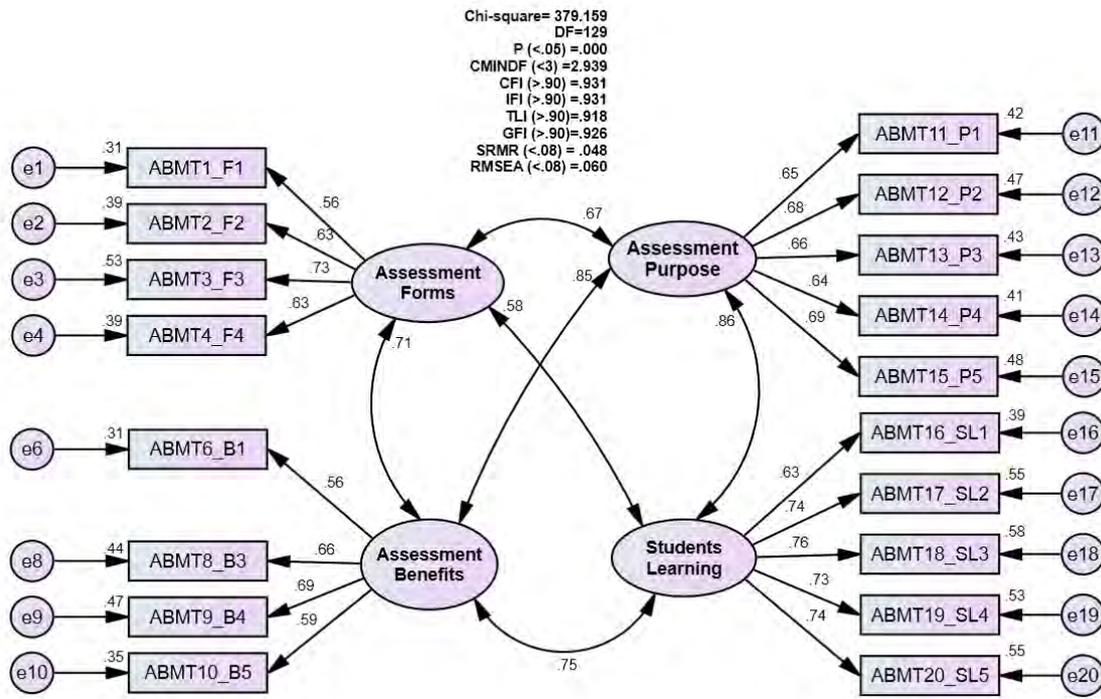


Figure 3. Confirmatory Factor Analysis for ABMT (Re-specified Model)

Additionally, the magnitudes of loadings for all items are statistically significant (T-value  $\geq 1.964$  and P-value  $\leq 0.05$ ) with the value of 0.50 as the acceptable coefficients Table 5 for exploratory level. CR for each construct in the hypothesized model met the acceptable criteria of 0.70 while the AVE did not meet the acceptable criteria (0.50). However, it can accomplish further analysis as long as the CR is achieved without meeting AVE (Fornell & Larcker, 1981a, 1981b), especially the issues of construct validity (Convergent Validity and Discriminant Validity) for the model of ABMT are beyond the objective of the present study.

Table 5. Results of Parameters of Measurement Model of ABMT

Construct	Items	B	SE	t-value	p	Factor loading	SMC	CR	AVE
Assessment Beliefs of Mathematical Thinking	Assessment Forms	1				0.703	0.495	0.92	0.74
	Assessment Benefits	1.22	0.15	8.029	0	0.974	0.948		
	Assessment Purpose	1.74	0.19	9.029	0	0.868	0.753		
	Students Learning	1.51	0.17	8.719	0	0.879	0.773		
Assessment Forms	ABMT1_F1	1				0.557	0.310	0.73	0.41
	ABMT2_F2	0.98	0.1	10.13	0	0.625	0.391		
	ABMT3_F3	1.26	0.12	10.96	0	0.731	0.535		
	ABMT4_F4	1.03	0.1	10.14	0	0.625	0.391		
Assessment Benefits	ABMT6_B1	1				0.555	0.308		
	ABMT8_B3	1.38	0.13	10.9	0	0.662	0.439	0.72	0.39
	ABMT9_B4	1.46	0.13	11.12	0	0.686	0.471		
	ABMT10_B5	0.98	0.1	10.2	0	0.595	0.354		

Table 5. Continued

Construct	Items	B	SE	t-value	p	Factor loading	SMC	CR	AVE
Assessment Purpose	ABMT11_P1	1				0.651	0.424	0.79	0.41
	ABMT12_P2	1	0.07	13.41	0	0.682	0.465		
	ABMT13_P3	1.01	0.08	13	0	0.657	0.431		
	ABMT14_P4	1	0.08	12.69	0	0.638	0.407		
	ABMT15_P5	1.06	0.08	13.57	0	0.692	0.479		
Students Learning	ABMT16_SL1	1				0.628	0.394	0.84	0.52
	ABMT17_SL2	1.32	0.1	13.89	0	0.742	0.55		
	ABMT18_SL3	1.32	0.09	14.15	0	0.762	0.58		
	ABMT19_SL4	1.13	0.08	13.7	0	0.728	0.53		
	ABMT20_SL5	1.17	0.08	13.88	0	0.741	0.549		

Key: B= Unstandardized Estimation, SE=Standard Error, P=Probability Value, SMC =Squared Multiple Regression, CR= Composite Reliability, AVE = Average Variance Extracted

Covariances Table 6 among the four constructs of the model of Assessment Beliefs of Mathematical Thinking are statistically significant (T-value  $\geq 1.964$  and P-value  $\leq 0.05$ ). Table 6 also shows that the value of coefficients of the relationship between every two dimensions of ABMT are positive and strong enough to be indicated that the ABMT model reaches a good construct.

Table 6. Results of Covariance among Constructs of ABMT Model

Construct	Construct	B	SE	t-value	P	r	SMC
Assessment Purpose	Students Learning	0.183	0.019	9.535	0.00	0.860	0.73
Assessment Forms	Students Learning	0.099	0.013	7.335	0.00	0.576	0.33
Assessment Forms	Assessment Benefits	0.100	0.013	7.456	0.00	0.710	0.50
Assessment Purpose	Assessment Forms	0.118	0.015	7.872	0.00	0.671	0.45
Assessment Benefits	Students Learning	0.129	0.016	8.278	0.00	0.753	0.56
Assessment Purpose	Assessment Benefits	0.148	0.017	8.694	0.00	0.845	0.71

Key: B= Unstandardized Estimation, SE=Standard Error, P=Probability Value, r=Correlation, SMC =Squared Multiple Regression

The results indicated in Table 6 revealed that there are positive relationships between every two dimensions of ABMT. Hence, the dimensions are linked together to form strong construction for teachers' assessment believes of mathematical thinking. Further, as it is shown in Table 7, the squared root of AVE for each construct of the hypothesized model was more than the SV except for a few relationships, establishing discriminant Validity for the model of ABMT.

Table 7. Results of Shared Variance (SV) and Squared Root of Average Variance Extracted (AVE)

Constructs	1	2	3	4
AP1 Assessment Purpose	<b>0.64</b>			
SL2 Students' Learning	0.73	<b>0.72</b>		
AF3 Assessment Forms	0.45	0.33	<b>0.64</b>	
AB4 Assessment Benefits	0.71	0.56	0.56	<b>0.63</b>

Bold size = Value of Squared Average Variance Extracted

As illustrated in Figure 4 and Table 4, the goodness fit indices for the second order CFA of ABMT, stated that its fit statistics are found within the acceptable criteria as Normed Chi-Squared CMINDF = 2.983 (achieved the threshold of  $< 3$ ), CFI = 0.928 (more than the threshold of  $> 0.90$ ), IFI = 0.916 (met the threshold of  $> 0.90$ ), TLI = 0.916 (reached the threshold of  $\geq 0.90$ ), GFI = 0.925 (accomplished the threshold of  $> 0.90$ ), SRMR = 0.050 (met the threshold of  $> 0.80$ ) and RMSEA = 0.061 (achieved the threshold of  $< 0.80$ ). Finally, the magnitudes of higher loadings for four sub-factors are statistically significant (T-Value  $\geq 1.964$  and P-Value  $\leq 0.05$ ).

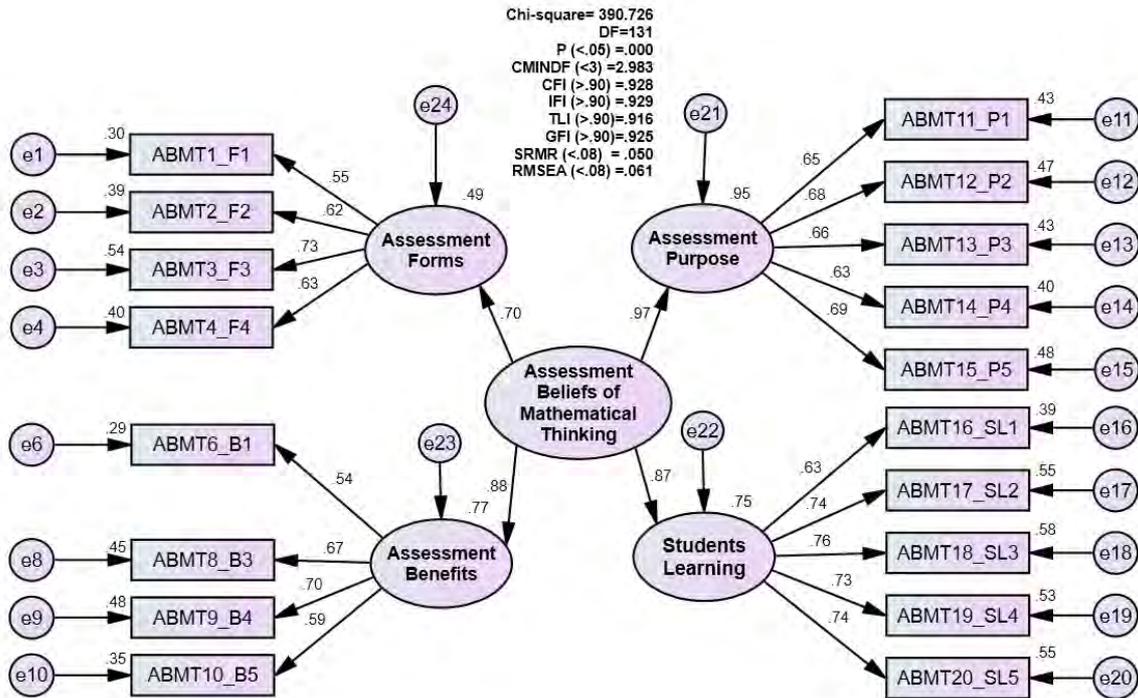


Figure 4. Second Order of Confirmatory Factor Analysis for ABMT

Consequently, the results of testing reliability for the instrument of assessment beliefs of mathematical thinking showed that all factors reached the accepted reliability level, and the model met the good fit indices of CFA. Therefore, the instrument of assessment beliefs of mathematical thinking can be used to measure teachers’ assessment beliefs of mathematical thinking.

### Discussion

This research tried to fill the gap of lacking instruments that can be used to examine the teachers’ assessment beliefs of mathematical thinking in the Arabic language, especially in Oman. Thus, this study developed an instrument of ABMT and tested it by using CFA. CFA was employed to validate the instrument to measure the mathematics teachers’ assessment beliefs of mathematical thinking. The instrument was established as a model consisting of four dimensions: assessment forms, assessment benefits, assessment purpose, and students’ learning. The original instrument contained 20 items, with 5 items for each dimension. After the testing of the scale using CFA, two items were deleted, one from the assessment forms and the other from the assessment benefits. The final instrument resulted in 18 items. The results of testing reliability revealed that all dimensions got the accepted reliability level, and the model reached the good fit indices of CFA. Therefore, the instrument of teachers’ assessment beliefs of mathematical thinking can be applied to assess teachers’ assessment beliefs of mathematical thinking.

The instrument of ABMT was developed through merging some previous instruments of assessment beliefs about mathematics in general with the National Council of Teachers of Mathematics (2000) standards of school mathematics. This produced a new scale that has new features for measuring teachers’ assessment beliefs of mathematical thinking. The instrument covers teachers’ assessment beliefs of mathematical thinking related to the assessment forms, assessment benefits, assessment purpose, and students’ learning. The results showed significant relationships among all factors. These results are consistent with Dandis (2013) which introduces an instrument of mathematics teachers’ assessment beliefs. The instrument was divided teachers’ assessment beliefs into six factors: assessment forms, assessment benefits, the connection between learning, teaching and assessment practices, students learning, assessment for learning, and assessment purpose. While the scale of Genc (2005) contained the teachers’ beliefs about diagnostic, summative, and formative assessments. Further, Büyükkarıcı (2014) instrument was developed to measure teachers’ beliefs about only formative assessment. Further, Muñoz et al. (2012) instrument focused on teachers’ beliefs about the responsibility of schools, the development of learning, and the teaching and documentation of students’ learning.

Moreover, the results showed that teachers believed assessment needs feedback to improve students’ confidence in their skills of mathematical thinking. Feedback could motivate students’ mathematical thinking and learning. It also can regulate the behaviours of the students, which leads to improving the students’ confidence while they learn mathematics (Smith, 2010). This result is supported by the previous studies, which indicated that students’ confidence could be improved if they are advised and assisted by their teachers. For example, Unal and Unal (2019) found that teachers

believed that assessment improves students' learning and thinking through the feedback given to the students. Yong and Sam (2008) stated that teachers' feedback could improve the students' mathematical thinking. Delice and Sevimli (2010) found that the confidence of the students in mathematics learning improved while they were assisted and encouraged by the teachers. Similarly, Nurlu (2015) found that students could achieve more confidence when they were supplied by their teachers' help.

The results also indicated that the teachers believed that addressing various assessment tools helps to meet the objectivity in assessing students' mathematical thinking. The teachers believed that a student's personality affects his/her responses to assessment tools. They also believed that the assessment provides them with the strengths and weaknesses in students' mathematical thinking skills. Assessments methods and tools help teachers to collect big information about students' learning of mathematics like understanding, abilities, and achievement (Ling et al., 2012; Tsamir & Tirosh, 2006). Siemon et al. (2017) resulted that teachers applied multiple assessment forms because they believed that addressing multiple assessment forms improve the students' understanding and reasoning of mathematics.

Furthermore, analyzing this information benefits teachers by determining the strengths and weaknesses of students' mathematics learning and thinking (Ali, 2015; Haynes et al., 2016). The finding of the analyzing can be used to concentrate on students' weaknesses to be improved. Teachers can also improve the strength points of students to continue their development in mathematical thinking. These results are in line with the results were found by Dandis (2013), who showed that mathematics teachers believed that assessment results supply them with information related to students' learning progress and the problems that students may face when they learn mathematics. Therefore, students with different tools will get more opportunities to present their mathematical understanding (Ali, 2015; Dandis, 2013). Reyes-Cedeno et al. (2019) showed that teachers believed that assessment practices help them to determine the starting point of improving the students' mathematical thinking.

### **Conclusion**

The instrument can be considered as an important instrument to be implemented for investigating the mathematics teachers' assessment beliefs of mathematical thinking in the Arab areas, especially in Oman either for public institutions or for self-assessment. It could introduce important information that can be used to improve the courses of preparing pre-service teachers to teach mathematics and evaluate students' mathematics learning. Further, the educational institutions can use this instrument to get information that helps to improve mathematics assessment in schools. This research introduced an instrument for mathematics researchers and educators to study the teachers' assessment beliefs of mathematical thinking from different aspects like the relation between teachers' assessment beliefs of mathematical thinking and some other variables relating to mathematics teaching such as assessment practices of mathematical thinking.

### **Recommendations**

The current research aimed to develop an instrument for teachers' assessment beliefs of mathematical thinking. The sample of the research belongs only to Omani public schools. The current instrument can be expanded to different locations and samples for achieving an enhanced measurement model. For instance, the instruments can be tested with different samples like teachers who teach in private schools or with mathematics teachers in another country. The instrument can be applied to test the relationship between teachers' assessment beliefs and practices of mathematical thinking. The research studied teachers' assessment beliefs of mathematical thinking by using a questionnaire, i.e. the quantitative methodology. Future research can study the same variable using both quantitative and qualitative methodologies. This may reveal more information concerning the research. Further researches can study the effect of teachers' experiences, anxiety, or attitudes toward mathematics on teachers' assessment beliefs of mathematical thinking.

The instrument can be considered as an important instrument to be implemented for investigating the mathematics teachers' assessment beliefs of mathematical thinking in the Arab areas, especially in Oman either for public institutions or for self-assessment. It could introduce important information that can be used to improve the courses of preparing pre-service teachers to teach mathematics and evaluate students' mathematics learning. Further, the educational institutions can use this instrument to get information that helps to improve mathematics assessment in schools. This research introduced an instrument for mathematics researchers and educators to study the teachers' assessment beliefs of mathematical thinking from different aspects like the relation between teachers' assessment beliefs of mathematical thinking and some other variables relating to mathematics teaching such as assessment practices of mathematical thinking.

### Limitation

It should be stressed that teachers who applied for the research instrument work in Omani public schools. All mathematics teachers in the public the Omani schools are Arab. Consequently, their views can be affected by the definitions of the terms used in Omani educational system because the assessment system in all public schools in Oman is similar, i.e. applying the same assessment documents published by the Ministry of Education. Those documents determine the conditions of each assessment tool and when it should be used. Another limitation of this study was that the teachers' assessment beliefs were gained in isolation from their conceptions of curriculum, teaching, and learning. This research also did not consider the correlation between variables such as age, teacher culture, and educational background with beliefs. The results of this study are limited to the research instruments because the research used only the self-reported questionnaire as the instruments.

### Authorship Contribution Statement

Alhunaini: Conceptualization, design, analysis, writing. Osman: editing/reviewing, supervision. Abdurab: Drafting manuscript, Statistical analysis

### References

- Abdullah, A., Fadil, S., Abd Rahman, S., Tahir, L., & Hamzah, M. (2019). Emerging patterns and problems of higher-order thinking skills (HOTS) mathematical problem-solving in the form-three assessment (PT3). *South African Journal of Education*, 39(2), 1-18. <https://doi.org/10.15700/saje.v39n2a1552>
- Ali, H. (2015). Toward differentiated assessment in a public college in Oman. *English Language Teaching*, 8(12), 27-36. <https://doi.org/10.5539/elt.v8n12p27>
- Bal, A., & Artut, P. (2019). Developing the mathematical thinking scale: Validity and reliability. *Journal of Çukurova University Faculty of Education*, 49(1), 278-315. <https://dergipark.org.tr/tr/download/article-file/1070196>
- Brown, G. (2011). Self-regulation of assessment beliefs and attitudes: a review of the students' conceptions of assessment inventory. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 31(6), 731-748. <https://doi.org/10.1080/01443410.2011.599836>
- Büyükkarcı, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction*, 7(1), 107-121. <https://dergipark.org.tr/en/download/article-file/59697>
- Calveric, S. (2010). *Elementary teachers' assessment beliefs and practices* [Unpublished doctoral dissertation]. Virginia Commonwealth University.
- Dandis, M. A. I. (2013). *Teachers' beliefs about assessment for learning: Introducing rubric in secondary education* [Unpublished doctoral dissertation]. Universidad de Granada.
- Delice, A., & Sevimli, E. (2010). An investigation of the pre-services teachers' ability of using multiple representations in problem-solving success: The case of definite integral. *Educational Sciences in Theory and Practice*, 10(1), 137-149. <https://files.eric.ed.gov/fulltext/EJ882726.pdf>
- Etheridge, B. (1989). *Mathematics: Grades 1 and 2 assessment* (ED312127). ERIC. <https://bit.ly/3bKl6Oa>
- Fornell, C., & Larcker, D. F. (1981a). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Fornell, C., & Larcker, D. F. (1981b). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research* 18(3), 382-388. <https://doi.org/10.1177/002224378101800313>
- Gebriil, A., & Brown, G. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 21(1), 16-33. <https://doi.org/10.1080/0969594X.2013.831030>
- Genc, E. (2005). *Development and validation of an instrument to evaluate science teachers' assessment beliefs and practices* [Unpublished doctoral dissertation]. Florida State University.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis*. Pearson Education Limited.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. *Long Range Planning*, 46, 1-12. <https://doi.org/10.1016/j.lrp.2013.01.001>
- Haynes, A., Lisic, E., Goltz, M., Stein, B., & Harris, K. (2016). Moving beyond assessment to improving students' critical thinking skills : A model for implementing change. *Journal of the Scholarship of Teaching and Learning*, 16(4), 44-61. <https://doi.org/10.14434/josotl.v16i4.19407>

- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. The Guilford Press.
- Ling, S. S., Lan, O. L., Suah, S. L., & Ong, S. L. (2012). Investigating assessment practices of in-service teachers. *International Online Journal of Educational Science*, 4(1), 91–106. <https://bit.ly/3mcfkem>
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *profile: Issues in teachers' professional development*, 14(1), 143-158. <https://bit.ly/3BKSI9e>
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. The National Council of Teachers of Mathematics, Inc.
- Nurlu, Ö. (2015). Investigation of teachers' mathematics teaching self-efficacy. *International Electronic Journal of Elementary Education*, 8(1), 489–508. <https://www.iejee.com/index.php/IEJEE/article/view/95/92>
- Privitera, G. J. (2020). *Research methods for the behavioral sciences*. SAGE Publications Ltd.
- Reyes-Cedeno, C., Rivas-Cun, H., Espinoza-Cevallos, C., & Rojas-Garcia, C. (2019). Assessment of the practices for early mathematics thinking in preschools of Pasaje City, Ecuador. *European Journal of Educational Research*, 8(4), 1063-1070. <https://doi.org/10.12973/eu-jer.8.4.1063>
- Siemon, D., Day, L., Stephens, M., Horne, M., Callingham, R., Watson, J., & Seah, R. (2017). Reframing mathematical futures: Using learning progressions to support mathematical thinking in the middle years. In A. Downton, S. Livy, & J. Hall (Eds.), *40 years on: We are still learning! Proceedings of the 40th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 647-650). MERGA.
- Smith, L. (2010). *The relationship of mathematics anxiety, mathematical beliefs, and instructional practices of elementary school teachers* [Unpublished doctoral dissertation]. The University of Southern Mississippi.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson Education, Inc.
- Tsamir, P., & Tirosh, D. (2006). Summing up and looking ahead: A personal perspective on infinite sets. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 49-63). Atelier Guimaec, s.r.o.
- Unal, A., & Unal, Z. (2019). An examination of K-12 teachers' assessment beliefs and practices in relation to years of teaching experience. *Georgia Educational Researcher*, 16(1), Article 2. <https://doi.org/10.20429/ger.2019.160102>
- Widiastuti, I. A., Mukminatien, N., Prayogo, J. A., & Irawati, E. (2020). Dissonances between teachers' beliefs and practices of formative assessment in EFL classes. *International Journal of Instruction*, 13(1), 71-84. <https://doi.org/10.29333/iji.2020.1315a>
- Yong, H. T., & Sam, L. C. (2008, August 28). *Implementing school-based assessment: The mathematical thinking assessment (MATA) framework* [Paper presentation]. Seminar Inovasi Pedagogi IPBL, Kuching, Sarawak, Malaysia.
- Youmans, A., Coombs, A., & Colgan, L. (2018). Early childhood educators' and teachers' early mathematics education knowledge, beliefs, and pedagogy. *Canadian Journal of Education*, 41(4), 1079-1104. <https://www.jstor.org/stable/26732972>
- Zainudin, M., Subali, B., & Jailani. (2018). Construct validity of mathematical creativity instrument: First-order and second-order confirmatory factor analysis. *International Journal of Instruction*, 12(3), 595-614. <https://bit.ly/3GWxitB>