www.ijemst.net

Three Foci of Mathematic Teachers' Considerations in Evaluation of School Students' Achievements

Ester Halfon
Shaanan College of Education, Israel

Yaniv Biton
Shaanan College of Education, Israel

To cite this article:

Halfon, E. & Biton, Y. (2022). Three foci of mathematic teachers' considerations in evaluation of school students' achievements. *International Journal of Education in Mathematics, Science, and Technology (IJEMST), 10*(1), 236-256. https://doi.org/10.46328/ijemst.1934

IJEMST

International Journal of Education in Mathematics, Science and Technology

2022, Vol. 10, No. 1, 236-256 https://doi.org/10.46328/ijemst.1934

Three Foci of Mathematic Teachers' Considerations in Evaluation of School Students' Achievements

Ester Halfon, Yaniv Biton

Article Info

Article History

Received:
12 July 2021
Accepted:
03 December 2021

Keywords

Assessment of mathematics academic achievements
Mathematic teachers' training
Teaching and assessment quality

Abstract

As part of efforts to improve the quality of mathematics' teaching and evaluation, we examined the focus of math teachers' considerations in evaluating students' achievements, as well as the links between these focuses, regarding differences between students and the validity and reliability of assessment methods and examinations. Based on the categorization of issues that concern math teachers while assessing their students, a self-report quantitative-questionnaire was constructed and validated. The questionnaire included three foci of consideration for mathematics evaluation: considerations regarding evaluating learners, considerations regarding the choice of assessment methods, and assessment considerations that are unique to mathematics; positive, high, and significant were found between them. The main conclusion is that math teachers aspire to evaluate their students on the basis of a broad picture, that considers the needs of students studying mathematics as a unique discipline, the available assessment methods and their ability to adapt them to a valid and reliable assessment in mathematics, and the overall need to focus on the challenges and difficulties which are unique to assessing the discipline of mathematics. The findings also strengthen the claim that assessing math learning has unique considerations that are distinguished from assessment considerations in other disciplines.

Introduction

Evaluating achievements is an integral part of teaching and understanding assessment considerations could contribute to improving the quality of teaching, and consequently, to improving student's achievements. This issue concerns many math teachers against the background of the advantages and disadvantages of qualitative and quantitative assessment methods in the realm of teaching mathematics (Silver & Mills, 2018). Most of the issues of concern to math teachers relate mainly to questions such as class size, student heterogeneity, the value of the thinking stages in the solution processes, and reliability and validity of math evaluation. Studies have also revealed that many teachers feel that they lack the knowledge of apply alternatives to the quantitative-traditional assessment (Abali et al., 2014; Al-Nouh et al., 2014; Biton & Halfon, in print; Black & Wiliam, 2012; Chiang, 2015; Darmody et al., 2020; Davison & Leung, 2009; Galustyan, 2017; Savickiene, 2011; Veldhuis & van den Heuvel-Panhuizen, 2020; Zhao et al., 2018). To date, the foci of mathematics teachers' considerations have not yet been thoroughly examined from the many different aspects of assessment which help them as they approach choosing and evaluating their students' achievements, knowledge and skills acquired in learning. Therefore, the purpose of the present study was to examine the relationships between the various considerations and identify and describe the foci of math teachers' considerations in evaluating their students' achievements, to try to answer questions of reliability and validity of ways of evaluating mathematics and their ability to reflect the mathematical knowledge and mathematical skills acquired by their students. The insight gained from the literature is that that a combination of quantitative assessment methods and a qualitative assessment method could enable math teachers obtain a more comprehensive, in-depth and accurate picture of their learners' achievements, knowledge and skills (Biton & Halfon, in print; National Council of Teachers of Mathematics, 2000; Nevo, 2002, 2006; Veldhuis & van den Heuvel-Panhuizen, 2020).

Two general approaches are common in mathematics assessment and evaluation: traditional assessment methods and alternative assessment methods. *Traditional assessment methods* are mainly based on quantitative tests aimed at diagnosing the knowledge and skills acquired by the students on the subject studied and quantifying their achievements in relation to the required level. This method is typically used as a summative evaluation and an examination of a final product (Pellegrino, 2003). *Alternative assessment methods* are based on the constructivist approach, thus enabling the teacher to examine the students' learning and research methods (Ültanir, 2012). Alternative assessment in mathematics is a powerful pedagogical tool in which learning is observed during its actual course. The use of alternative assessment if based on active participation in the assessment processes, alongside the learners' learning processes and adapting the learning content to their abilities and needs (Silver & Mills, 2018).

Alternative assessment methods in mathematics are diverse: *descriptive assessment*, which includes open questions or a requirement that students detail the problem-solving process, so that teachers can analyze the way students resolved (and not just the end result), and accordingly – help them understand their learning and improve it (Cai et al., 2020c; Kim & Noh, 2010; National Council of Teachers of Mathematics, 2000). Other alternative assessment methods in math are oral exams and interviews (Kulm, 2013; Watt, 2005; Zhao et al., 2018). Another alternative evaluation method is the creation of a "concept map," through which students learn to identify the contexts and interactions between the mathematical topics they had learned, and a representational understanding of the idea studied. Another effective alternative assessment approach is peer assessment, which helps develop meta-cognitive thinking and increases learners' self-awareness of their strengths and weaknesses in learning mathematics. By writing activity reports or journal writing, learners are required to identify what the knowledge is that they are learning, and the context of the knowledge to previous knowledge they had acquired. With these methods the portfolio can also be mentioned. The portfolio includes works and documents attesting to research and learning, and which develop reflective and creative thinking. Alternative assessment by observation enables teachers to learn about the types of interactions and strategies and follow their students' learning processes. (Cai et la., 2020a; Shahbari et al., 2018; Silver & Mills, 2018).

In the recent decades, it is agreed that planning teaching should be based on diverse data (such as test scores, performing exercises, in-class assignments and tasks, the correctness of ways of solving and the degree to which students use strategies while performing tasks, the products of classroom discourse, and students' different responses). Such data that may inform the teacher about the way students think is particularly helpful in allowing teachers to anticipate students' responses to educational tasks. Thus, help teachers make better decisions in planning teaching units and implementing them, and improving the quality of their teaching (Cai et al., 2020a, 2020c). In Israel, the internal and external *Meitzav* tests (school efficiency and growth measures) were instituted in the fifth and eighth grades (Beller, 2012, 2013), and they provide data on the level of students, class, and school, which constitute formative evaluation and contribute to improving teaching quality.

Data-based teaching/learning opportunities are based on interactions between three elements: math tasks, teaching methods, and students. The nature of the interactions between the three elements will determine whether the learning experience will become a learning opportunity for the students (Cai et al., 2017, 2020 a).
The data that emerge from an alternative assessment can serve as a basis for describing the thought processes of students during the experiences and for improving teaching methods and teaching practices. These data are include knowledge of the way students in a certain class could respond to certain aspects of teaching tasks, in addition to knowledge about patterns observed in all classes (Cai et al., 2020b). These data are used as tool to collect, process, and analyze data on individual students, increasing teachers' understanding of students' mathematical learning experiences (Cai et al., 2018). Additionally, these data can be useful also because the strategy that students used for solving problems exposed their thought procedures to reach to solution.
Documenting the strategies that the students used can help teachers predict how students can understand new problems introduced during the class and how they will think about them. Such data also serve as a basis for pedagogical professional knowledge, as a part of improving the long-term teaching of mathematics (Cai et al., 2020b), helping to promote teachers' research on ways to improve teaching and promote student learning and achievements (Cai et al., 2018; National Council of Teachers of Mathematics, 2013). Although the alternative-assessment tasks are complex and require a great investment and much time, there is much importance in defining the content domain that is being assessed, preparing indicators to cover this content domain. It is also important that the content represents the latest understanding of the field, and therefore the place of experts is important to approve the tasks correctly in terms of their context, meaning, and value. To achieve the goal of assessment, one must demand evidence that carrying out the assignments is not a matter of learning by rote, but one that represents the entire learning process that the student underwent to fulfill the assignment (Schiefer et al., 2019).

Alternative assessment methods are correlated with higher achievement levels, learning motivation and diligence, and students' perception of student efficacy (Sahin & Abali Ozturk, 2014). Hence, traditional evaluation methods should be used with alternative ones (Chiang, 2015; Veldhuis & van den Heuvel-Panhuizen, 2020).
The advantages and strengths of the alternative mathematics assessment method, alongside the difficulties and challenges that characterize it, have led many researchers to investigate the subject (Ediger, 2013; Kulm, 2013; Watt, 2005; Zhao et al., 2018). On the one hand, the studies indicate that alternative assessment methods help

improve the assimilation of learners' learning processes, improve their academic achievement, develop personal learning potential, and improve their positive attitude toward mathematics (Davison & Leung, 2009; Ediger, 2013; Galustyan, 2017; Kulm, 2013; Sahin & Abali Ozturk, 2014; Savickiene, 2011; Veldhuis & van den Heuvel-Panhuizen, 2020; Zhao et al., 2018). On the other hand, despite the increasing recognition of alternative assessment methods as contributing to the quality of learning of learners and promoting their achievements, there are difficulties in their application (Briggset al., 2012; Cai et al., 2020c; Kingston & Nash, 2011, 2012), and in educating teachers to use informed considerations to choose valid and reliable assessment methods. Perusing the research literature revealed that math teachers have difficulties with the task planning and transmitting stages.

Furthermore, they are concerned about the level of validity and reliability of alternative assessment methods, and also report difficulties in recruiting resources to carry out the alternative assessment that requires higher financial investment than a quantitative exam (Davison & Leung, 2009; Kulm, 2013; Levy-Vered & Nasser-Abu Alhija, 2015; Li et al., 2019; Savickiene, 2011; Veldhuis & van den Heuvel-Panhuizen, 2014, 2020; Watt, 2005). Teachers have different approaches towards alternative assessment; some are familiar with a variety of alternative assessment methods and use them equally, being aware of their contribution to instruction and learning. Others use them less than average or very little (Cai et al., 2020c; Veldhuis & van den Heuvel-Panhuizen, 2014; Zhao et al., 2018). The limited use of alternative assessment affected by its negative perception, teachers' belief system, and lack of knowledge on the subject of alternative assessment (Kulm, 2013; Levy-Vered & Nasser-Abu Alhija, 2015).

Providing guidance and support to teachers who develop and use alternative assessment in mathematics, could contribute to improving students' achievement in mathematics (Mandinach & Jackson, 2012; Silver & Mills, 2018; Veldhuis & van den Heuvel-Panhuizen, 2020). In previous qualitative research, Biton and Halfon (in print) identified three key areas (foci) of mathematics assessment that concern math teachers and student teachers: the validity and reliability of math tests, the heterogeneity of the evaluated students, and the students' level of knowledge and achievements as indicated by their assessment. Following, this quantitative study examined whether, and to what degree are math teachers concerned with these considerations in each of the identified foci? To what degree do these considerations contribute to ensuring that their assessments are reliable, valid, and reflect the knowledge and skills acquired by their students?

Methodology

Aim

The aim of the study is to examine the connections between the main foci in math teachers' considerations regarding the assessment of their students' achievements in mathematics, against the background of the variance of students on the one hand, and the validity and reliability of the tests and assessments on the other. Another aim is to identify and describe the foci of math teachers' considerations in evaluating their students' achievements, to try and respond to questions of the reliability and validity of ways of assessment for mathematics, and the degree to which they can reflect the knowledge and mathematical skills acquired by their students. These aims have high importance as part of the effort to improve teaching and evaluation in general, and in mathematics particularly.

Research Questions

What are the correlations between the three foci of considerations of Mathematics teachers' evaluation of their students' achievements?

To what degree do the considerations in these three foci contribute to confirm that their evaluations are reliable, valid and reflect students achievements?

Are there differences in considerations between math teachers according to their background in evaluation, specialization in mathematics, number of classes the teacher teaches, and the number of students the teacher teaches?

Regarding research question 1, it is hypothesized that there are positive and strong relationships between the three foci of math teachers' considerations in evaluating their students' achievements. No hypothesis were formulated for Research question 2 and 3 because this study is the first of its kind.

Research Design

The study was conducted using a quantitative approach. The research instrument was a self-report questionnaire that was constructed and validated on the basis of the findings of a previous qualitative research (Biton & Halfon, in print).

Procedure

The research design is quantitative – a questionnaire was distributed to math teachers. Based on the concerns of math teachers that were found in a previous study (Biton & Halfon, in print), concerns were selected according to the following criteria: relevance to evaluation in mathematics and quality and clarity of the phrasing. After the questionnaire was constructed, it was transferred for validation to expert judges who evaluated – separately and together – how much each statement belongs, in terms of content and wording, to the group of statements to which it was associated. The questionnaire was sent to various groups of teachers and student teachers (a total of 140 teachers), who were asked to answer it anonymously. A total of 84 responses to the questionnaire were received (response rate of 60%).

Population and Sample

Background information of the math teachers who responded to the questionnaire is presented in Table 1. About half of the teachers had over ten years of seniority in teaching generally and mathematics specifically, a third had between four and nine years of seniority and the rest (10.7%) three years or less. Most of them (79.8%) have specialized in Mathematics, and 78.6% have some background in student achievements' Evaluation and assessment.

Table 1. Mathematic Teachers' Background

		N	%
Years of seniority in teaching	1-3	9	10.7
	4-9	29	34.5
	+10	46	54.8
Years of seniority teaching mathematic	1-3	12	14.3
	4-9	32	38.1
	+10	40	47.6
Specialization in teaching mathematic	Yes	67	79.8
	No	17	20.2
Background in student achievements' Evaluation and assessment	Academic/advance course	66	78.6
	None	18	21.4

***N* = 84**

The professional background characteristics of the math teachers are presented in Table 2.

Table 2. Mathematic Teachers' Work Characteristics

		N	%
School level	Elementary school	77	91.7
	Middle school	7	8.3
Number of classes the teacher teaches	one class	18	21.4
	2-4 classes	59	70.2
	5+ classes	5	6.0
	(no response)	2	2.4
Number of students the teacher teaches	< 20 students	11	13.1
	20-30 students	36	42.9
	30 students or more	35	41.7
	(no response)	2	2.4

***N* = 84**

Most of the teachers work in an Elementary school (91.7%) and the rest in middle school. Most of them teach two classes or more (76.2%) or one class (21.4%). Most of them teach 20 students or more (82.6%); 13.1% reported that they teach less than 20 students.

Tool

Based on the first qualitative stage of the research (Biton & Halfon, in print), statements were produced from mathematic teachers' consideration regarding evaluating their students' knowledge and competencies. These teachers participated in three in-service courses for math teachers and two academic courses for student-teachers on evaluating achievement in mathematics. Following a systematic content analysis of their answers, a 25-statement questionnaire was constructed.

The statements describing considerations were categorized into three foci: Ten considerations focused on assessment of students, based on their abilities, difficulties, and variance ($\alpha = .80$), six considerations focused on methods of assessment of learning mathematics which are available to the teacher ($\alpha = .78$), nine considerations focused on assessment methods in mathematics as a discipline *sui generis* ($\alpha = .73$). The instruction given the teachers was: "Following are presented 25 concerns. Please rank the extent to which each consideration concerns you whilst evaluating your students' knowledge and achievements in learning mathematics." The respondents were asked to rank each statement on a 5-point Likert scale (from 1 = *not at all*, to 5 = *very much*).

The three foci of the statements and their distribution in the questionnaire are presented in Figure 1.

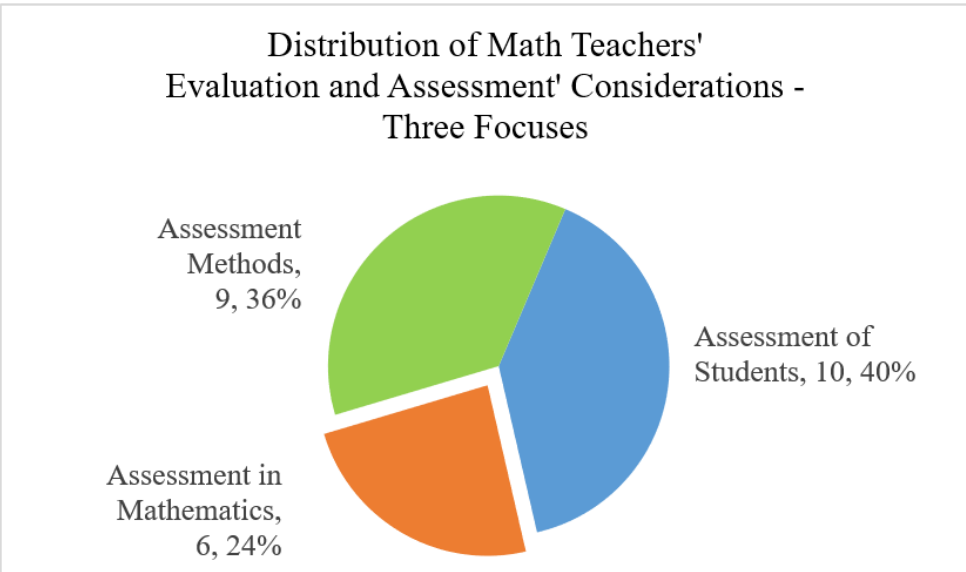

Figure 1. Three Foci of Math Teachers' Evaluation Considerations

Data Analysis

Descriptive statistics included frequencies and percentages, averages and standard deviations. The correlations between the three foci of consideration were calculated using Pearson r correlation coefficients. To compare math teachers by different characteristics, t tests were calculated between two independent groups, and a one-way analysis of variance was conducted between three groups.

Findings

Correlations between the Three Foci of Math Teachers Evaluating Consideration

In response to first research question, positive, high, and significant correlations were found among the three foci of consideration. For instance, teachers tend to focus simultaneously on all three foci together; the more they focus on adapting the assessment to the learners, their abilities, and difficulties, the more they focus on considerations in choosing ways of evaluation (r = .78, $p \leq .000$) in general and mathematics in particular (r = .71, $p \leq .000$). In other words, the percent of explained variance of both these foci of consideration (focusing on learners and focusing on assessment methods) is 50.4% (the variance explained by a square of the correlation); the remaining variance (49.6%) is explained by other factors.

The correlation between using considerations focused on assessment methods in general, and in mathematics in particular is positive, relatively high and significant (r = .64, $p \leq .000$). In other words, the percent of explained variance of both these foci of consideration is 41% (the variance explained by a square of the correlation); the remaining variance (59%) is explained by other factors. In summary, positive, strong, and significant links were found between the degree to which math teachers exercise considerations in the three foci when evaluating their students.

Math Teachers' Considerations in the Three Foci

In response to the second research question, the following tables present means and standard deviations of the three foci of consideration (evaluating learners, ways of assessment, and assessment in mathematics). Additionally, the averages and standard deviations of the statements are presented in the tables, as they were included in each of the three foci of consideration, and for each of the statements the frequency and percent of respondents who rated it as a consideration which is taken into account to a great degree /very much, somewhat, or very little/not at all.

Considerations Focused on Learners' Evaluation and Assessment

The math teachers' considerations focused on evaluating the learners are presented in Table 3 and Figure 2.

Table 3. Considerations Focused on Learners' Evaluation and Assessment

The statement		Considered		Somewhat considered		Not considered			
		f	*%*	*f*	*%*	*f*	*%*	*M*	*SD*
Average - Considerations in evaluating learners								**3.5**	**0.68**
19	No breakdown of the way of thinking – only a final answer. Students struggle to explain how they reached the result.	61	72.6	18	21.4	5	6.0	3.9	0.90

	The statement	Considered *f*	Considered %	Somewhat considered *f*	Somewhat considered %	Not considered *f*	Not considered %	*M*	*SD*
21	Difficulty of the examinee in understanding the formulation of a test question (reading comprehension)	56	66.7	20	23.8	8	9.5	3.9	1.05
3	The student understands the material in the classroom fails the test or receives a low grade that is at a relative gap to the knowledge	55	65.5	17	20.2	12	14.3	3.8	1.10
1	Children who become anxious and during the knowledge is not truly expressed	49	58.3	25	29.8	10	11.9	3.8	1.02
18	A teacher's difficulty to evaluate achievements against desire, effort, and ability: if the child tries hard and makes an effort but does not necessarily manage to reach score of 100. If he gets 70, he'll see it as a failure	50	59.5	22	26.2	12	14.3	3.6	1.04
5	How do you know that the student worked alone and wasn't helped by anyone else?	36	42.9	20	23.8	28	33.3	3.2	1.13
12	Each student is different in terms of knowledge, level, strengths, background, and therefore assessing achievement is not something that is certain and generalizable for the entire class	35	41.7	24	28.6	25	29.8	3.2	1.25
20	Subjective assessment of the teacher (influenced by previous acquaintance with the student)	36	42.9	19	22.6	29	34.5	3.1	1.30
10	Assessment does not check student's personal progress	30	35.7	27	32.1	27	32.1	3.1	1.17

	The statement	Considered		Somewhat considered		Not considered			
		f	%	*f*	%	*f*	%	*M*	*SD*
14	It's hard to evaluate a student with unclear handwriting	32	38.1	22	26.2	30	35.7	3.0	1.25

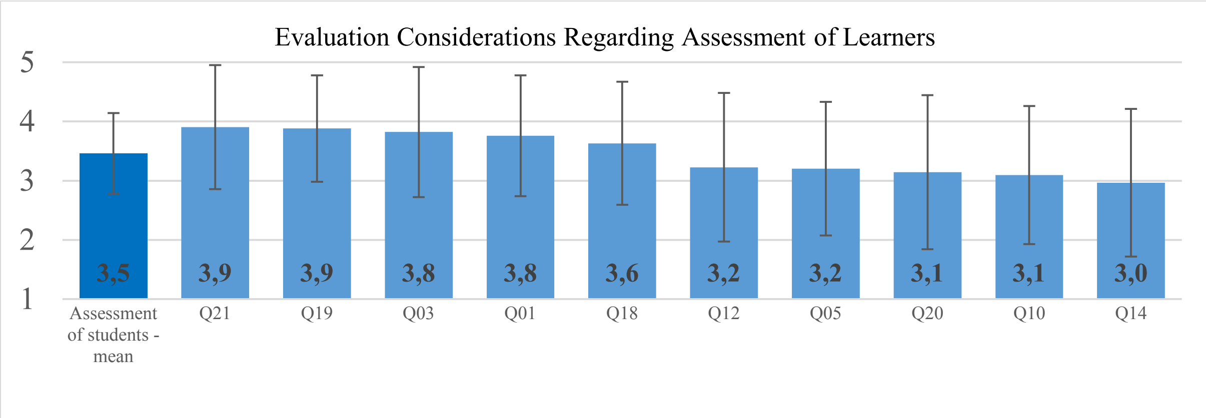

Figure 2. Evaluation Considerations focused on Assessment of Learners (Means and Standard Deviations)

On average, teachers are focused on considerations related to evaluating learners to a moderate or a large degree (M = 3.5, SD = 0.68). The two considerations related to evaluating learners in which math teachers are most focused while evaluating their students are: (21) The examinee's difficulty in understanding the formulation of a test question (reading comprehension) (M = 3.9, SD = 1.05) and (19) No breakdown of the way of thinking – only a final answer. Students struggle to explain how they reached the result. (M = 3.9, SD = 0.90). To the least degree (somewhat), among the considerations related to evaluating learners, math teachers focus on three considerations: (14) It's hard to evaluate a student with unclear handwriting (M = 3.0, SD = 1.25), (10) Assessment does not check student's personal progress (M = 3.1, SD = 1.17), and (20) Subjective assessment of the teacher (influenced by previous acquaintance with the student) (mean= 3.1, SD = 1.30). Significant differences were found in statement (12) regarding the claim that each student is different regarding knowledge, level, strengths and background evaluation and thus, assessment of the entire class is uncertain and not generalizable. Teachers who do not specialized in teaching mathematic agreed (M = 3.09 , SD = 1.23) that each student is different regarding knowledge, level, strengths and background evaluation and thus, assessment of the entire class is uncertain and not generalizable, significantly less ($t_{(82)}$ = 2.02, p = .047) than teachers who specialized in teaching mathematic (M = 3.76 , SD = 1.25). Additionally, teacher who teach less than 20 students (M = 3.36 , SD = 1.63) or 30 students or more (M = 3.54, SD = 1.25) agreed to a certain extent with this statement, but significantly more ($F_{(2,79)}$ = 3.83, p = .039) than teachers who teach medium classes (20-30 students) (M = 2.81, SD = 1.04). However, regarding the differences between mathematics teachers according to various background characteristics, on average no significant differences were found in these considerations (M = and each of the statements) according to mathematic teachers' background (years of seniority in teaching, years of seniority teaching mathematic, specialization in teaching mathematic, background in student achievements' evaluation and assessment) and mathematic teachers' work characteristics (school level, number of classes the teacher teaches).

Considerations Focused on Evaluation and Assessment Methods

The teacher's considerations focused on evaluation and assessment methods are presented in Table 4 and Figure 3.

Table 4. Considerations Focused on Evaluation and Assessment Methods

	Statement	Considered		Somewhat considered		Not considered			
		f	*%*	*F*	*%*	*F*	*%*	*M*	*SD*
Average - Considerations focused on the assessment methods								**3.6**	**0.63**
25	Do the Meitzav exams (school efficiency and growth indices), which the students are intentionally prepared to, really provide a reliable picture? Do the "Meitzav" exams (school efficiency and growth indices), which the students are intentionally prepared to, really provide a reliable picture?	66	78.6	12	14.3	6	7.1	4.2	1.08
22	Difficult to pay personal attention in a large group of students (also medium group)	64	76.2	17	20.2	3	3.6	4.2	0.87
24	Mastery varies – while studying – the students know it, but when the content is not used – they forget it	57	67.9	20	23.8	7	8.3	3.9	0.95
23	Difficulty in checking homework	46	54.8	23	27.4	15	17.9	3.6	1.18
4	The gap between the report card which is based on a quantitative score and an assessment that is not necessarily based on a quantitative score (including an alternative assessment)	46	54.8	21	25.0	17	20.2	3.5	1.10

Statement		Considered		Somewhat considered		Not considered			
		f	%	*F*	%	*F*	%	*M*	*SD*
2	How to evaluate a student if not on an exam? How can you evaluate when you give a thinking task in pairs or a group?	40	47.6	25	29.8	19	22.6	3.4	1.20
16	In the test: equal scoring for each question despite differences in the level of difficulty (the expectation: a more difficult question will receive a high score)	40	47.6	22	26.2	22	26.2	3.3	1.16
13	The questions are not in accordance with what was learned in class	36	42.9	21	25.0	27	32.1	3.1	1.30
8	How much do you take off for a recurring error	28	33.3	25	29.8	31	36.9	2.9	1.24

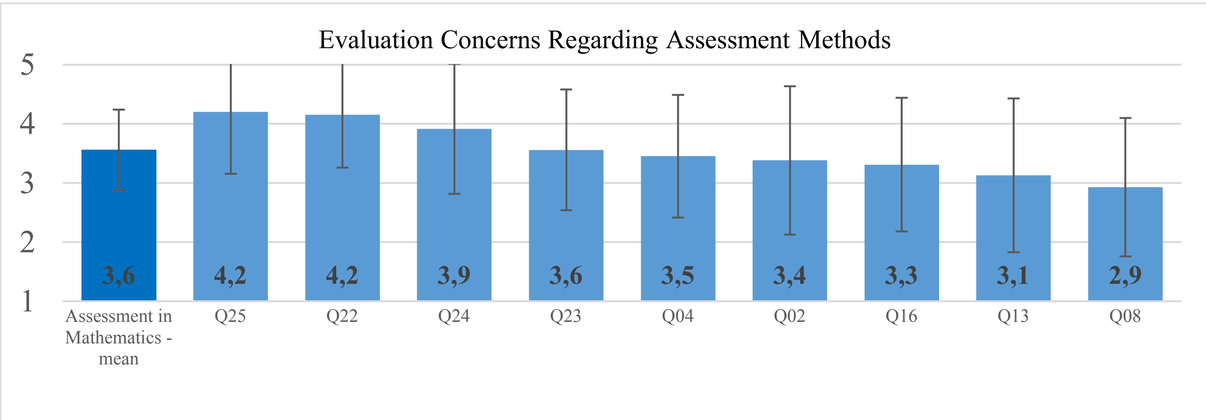

Figure 3. Evaluation Concerns regarding Assessment Methods (Means and Standard Deviations)

On average, teachers are focused on considerations related assessment methods to a moderate to high degree (*M* = 3.6, *SD* = 0.63). The three considerations related to assessment methods on which math teachers are most focused while evaluating their students are: (25) Do the "Meitzav" exams (school efficiency and growth indices), which the students are intentionally prepared to, really provide a reliable picture? (*M* = 4.2, *SD* = 1.08), (22) Difficult to pay personal attention in a large group of students (also medium group) (mean= 4.2, SD = 0.87), and (24) Mastery varies – while studying – the students know it, but when the content is not used – they forget it (*M* = 3.9, *SD* = 0.95). The consideration (8) How much do you take off for a recurring error is only moderately at the focus of consideration (*M* = 2.9, *SD* = 1.24).

In terms of the differences between teachers and mathematics according to different characteristics, on average no significant differences were found in these considerations (M = and each of the statements) according to *mathematic teachers' background* (years of seniority in teaching, years of seniority teaching mathematic, specialization in teaching mathematic, background in student achievements' evaluation and assessment) and *mathematic teachers' work characteristics* (school level, number of classes the teacher teaches). However, Significant differences were found in statement (12) regarding the claim that each student is different regarding knowledge, level, strengths and background evaluation and thus, assessment of the entire class is indefinite nor generalizable.

Teachers who specialized in teaching mathematic agreed that each student is different regarding knowledge, level, strengths and background evaluation and thus, assessment of the entire class is indefinite nor generalizable, significantly more ($t_{(82)}$ = 2.02, p = .047) than teachers who did not specialized in teaching mathematic, who agreed with this statement only to a certain degree. Additionally, teacher who teach less than 20 students (M = 3.36 , SD = 1.63) or 30 students or more (M = 3.54, SD = 1.25) agreed to a certain degree with this statement, but significantly more ($F_{(2,79)}$ = 3.83, p = .039) than teachers who teach medium classes (20-30 students) (M = 2.81 , SD = 1.04) who tend not less agree with this statement.

No significant differences were found in these considerations (M = and each of the statements) according to *mathematic teachers' background* (years of seniority in teaching, years of seniority teaching mathematic, specialization in teaching mathematic) and *mathematic teachers' work characteristics* (school level, number of classes the teacher teaches, number of students the teacher teaches). Additionally, teachers who have a background in student achievements' evaluation and assessment agreed that (22) it is difficult to pay personal attention in a large or medium group of students (M = 4.06, SD = .88), and to a certain extent that they focus on (23) the difficulty in checking homework (M = 3.42, SD = 1.20) - significantly ($t_{(82)}$ = -2.05, p = .049, $t_{(82)}$ = -2.06, p = .043 respectively) **less** than teachers with no previous background in student achievements' evaluation and assessment (M = 4.50, SD = .79, mean 4.06, SD = 0.94 respectively). Regarding considerations focused on evaluation and assessment methods, teachers who teach five classes or more (M = 4.20, SD = .84) take under consideration to a high extent that (24) Mastery varies – while studying – the students know it, but when the content is not used – they tend to forget it, more than teacher who teach significantly less ($F_{(2,79)}$ = 3.77, p = .027) than four classes or less.

It was also found that teachers who teach 30 students or more focus their considerations to a certain-high degree (M = 3.51, SD = 1.22) on (13) The questions are not in accordance with what was learned in the class, significantly more ($F_{(2,79)}$ = 4.57, p = .013) that teachers who teach 20-30 students (M = 2.64, SD = 1.20) or less than 20 students (M = 3.18, SD = 1.33). Teachers who teach 30 students or more focus their considerations also focus their considerations to a very high degree (M = 4.60, SD = .60) in (25) The reliability and validity of the of the Meitzav exams – significantly more ($F_{(2,79)}$ = 4.32, p = .017) than teacher who teach 20-30 students (M = 3.92, SD = 1.18) or less than 20 students (M = 4.09, SD = 1.30).

Considerations Focused on Evaluation and Assessment in Mathematics

The unique considerations in the math assessment are presented in Table 5 and Figure 4.

Table 5. Considerations Focused on Evaluation and Assessment in Mathematics

Statement		Considered		Somewhat considered		Not considered			
		f	%	*f*	%	*f*	%	*M*	*SD*
Average - Considerations focused on math evaluation								**3.4**	**0.78**
6	Difficulty in seeing the thought process that led to a solution (whether an mistake or a correct answer), hence the problem of how to handle difficulty / error	48	57.1	27	32.1	9	10.7	3.7	0.99
15	Evaluation of partial / full work with a mistake	28	33.3	29	34.5	27	32.1	3.6	1.13
9	Evaluation on way versus evaluation on result	40	47.6	28	33.3	16	19.1	3.4	1.18
7	Numerical grade does not allow the student to correct and improve himself, as s/he does not know or does not understand his/her errors	37	44.1	27	32.1	20	23.8	3.3	1.24
17	Usually, exams check the result and not the process. Sometimes a incorrect answer is rejected although the line of though was correct	48	57.1	22	26.2	14	16.7	3.2	1.09
11	Deliberation in scoring a solution that is not fully written, yet it is clear that the student understands and knows the solution	37	44.1	26	31.0	21	25.0	3.0	1.10

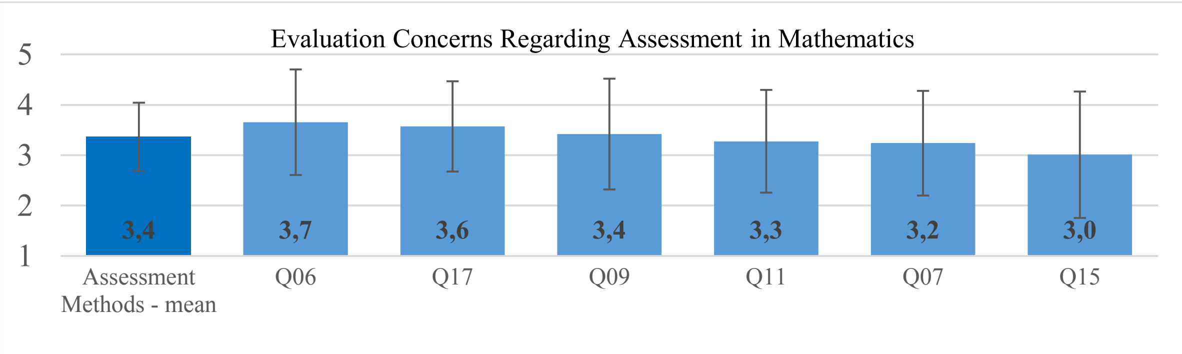

Figure 4. Evaluation Concerns regarding Assessment in Mathematics (Means and Standard Deviations)

On average, teachers are focused on considerations related to the assessment methods in mathematics to a moderate degree (*M* = 3.4, *SD* = 0.78). The considerations associated with the assessment methods in mathemati**cs** in the highest way – moderate to very high, focus on (6) Difficulty in seeing the thought process that led to a solution (whether a mistake or a correct answer), hence the problem of how to handle difficulty / error (mean= 3.7, *SD* = 0.99). A moderate degree (the lowest in this category of considerations) is the consideration focused (15) on evaluation of partial / full work but with a moderate mistake (*M* = 3.0, *SD* = 1.10).

As for the differences between math teachers according to different characteristics, on average, no significant differences were found in these considerations (*M* = and each of the statements) according to *mathematic teachers' background* (years of seniority in teaching, years of seniority teaching mathematic, specialization in teaching mathematic, background in student achievements' evaluation and assessment) and *mathematic teachers' work characteristics* (school level, number of classes the teacher teaches).However, teachers who teach less than 20 students focus to a high extent (*M* = 4.09, *SD* = 1.14) on (17) that consideration that usually, exams check results and not process (sometimes an incorrect answer is rejected although the line of though was correct) – significantly ($F_{(2,79)}$ = 4.89, *p* = .010) more than teachers who teach 20-30 students (*M* = 3.14, *SD* = 1.15) or 30 students or more (*M* = 3.80, *SD* = .99).

Discussion

This study examined the focus of the considerations of math teachers in assessing the achievements of their students, against the background of student variance and the validity and reliability of the tests – as part of the efforts to improve the quality of teaching and evaluation in general and in mathematics in particular. Three foci of considerations of assessing math learners were presented to teachers in a questionnaire: assessing learners based on their abilities, difficulties, and variance; the various assessment methods available to teachers; and the assessment methods that are unique to mathematics. The finding reveal that the correlations between the three foci of consideration are positive, high, and significant; the more math teachers focus on adapting the assessment to the specific learners they teach, the more they take into account their abilities and difficulties – the more they also focus on considerations related to their choice of assessment methods in general and mathematics in particular (Biton & Halfon, in print).

Thus, teachers who use considerations in evaluating students, simultaneously apply considerations relating to appropriate ways of evaluation in general and in mathematics in particular – and vice versa. Therefore, as a means of improving the cycle of teaching, learning, and evaluation (Birnbaum et al., 2006) – it is important to encourage teachers to expand their knowledge and skills in ways of assessment that are suitable for mathematics as a distinct discipline on the one hand, and to direct them toward awareness of the abilities and difficulties of students in assessment situations that may affect the reliability of the assessment and its results. Although the correlations that were found are relatively high, together, they explain only about half of their common variance in the foci of their evaluation in mathematics considerations. The remaining variance is explained by factors not examined in the present study and could deepen the understanding regarding other consideration that teachers use when they assess their students. These factors could be related to the environmental physcial conditions where the teaching takes place (number of students in class, noisy or quiet environment, lighting, test hours, etc.), the resources available to teachers (smart boards, computers and peripherals, and specialized software for teaching and learning mathematics), the teachers' education, quality and professional level in mathematics, etc. (see, for example, Abali Öztürk & Şahin, 2014; Biton & Halfon, in print; Darmody et al., 2020; Veldhuis & van den Heuvel-Panhuizen, 2020). In follow-up studies, it is proposed to find out what these additional factors are as well as other factors that could affect the relationship between considerations focused on assessment methods in general and mathematics in particular , and how they can be improved in order to increase the reliability of assessment in mathematics and its contribution to the teaching/learning cycle.

Among the considerations related to evaluating learners on which math teachers are focused while evaluating their students, the teachers focus mainly on the possible difficulty of the examinees in understanding the formulation of the questions presented to them in the test (reading comprehension: familiarity with the formal language of mathematics, the meaning of mathematical expressions whose formulation in relation to mathematics does not necessarily overlap with their daily meaning, as well as the great importance of noting the details and being meticulous about accuracy), and the need to grade an answer in which the way of thinking and the process of solving were not detailed, only the final answer written, due to students' difficulty to explain how they reached the result. These two foci of consideration characterize, to a great degree, the discipline of mathematics and should be taken into account when planning and conducting the evaluation. It is possible to respond to this variety of difficulties by assessing the learners using a combination of alternative assessment alongside the quantitative exams (Silver & Mills, 2018), as the advantages of qualitative assessment is the possibility it gives teachers to verify by providing feedback even in the learning stages (and not just feedback on the final product) that the students understood the questions and assignments, because they know and have mastered the formal language of mathematics and its unique expressions, they have formulated their work with the proper meticulousness and precision, and that the ways to solve the problems presented to them and completing the assignments are clear and appropriate, namely, to receive a comprehensive, in-depth, and accurate picture (Cai et al., 2020c. Kim & Noh, 2010; Kulm, 2013; Shahbari et al., 2018; Silver & Mills, 2018; Veldhuis & van den Heuvel-Panhuizen, 2020). However, as has also emerged in the literature, teachers do not have enough knowledge of planning executing alternatives that will complement the disadvantages of quantitative assessment (Al-Nouh et al., 2014; Black & William, 2012; Zhao et al., 2018). Therefore, it is appropriate that teachers' in-service learning will integrate, side by side, ways of evaluation in both methods, so that teachers will be clear about how each method complements the disadvantages of the other. In this context, the issue of data-based evaluation arises, both quantitative and qualitative data, which allow teachers to expose their students' manner of thinking (Cai et al., 2020c), thus improving their own quality of teaching (Cai et al., 2017).

Thus, among the considerations related to way of evaluation**,** the teachers focus first and foremost on the reliability of external tests (such as the Meitzav) in which a great deal of work is invested a lot of work in preparing for them, sometimes – at the expense of deepening the understanding and teaching of procedures for solving problems of various types, difficulty in paying personal attention to individual students when learning is conducted in a large and medium-size groups, and when students' retention and mastery of the material can change because of the frequency of dealing with each subject. These three considerations reflect varied aspects of mathematics evaluation: the aspect of external evaluation (the Meitzav tests), the aspect of the student as an individual, and the aspect of learning and retention. The reciprocal relations between these three aspects could contribute to turning the learning experience into an opportunity for effective learning for the students (Cai et al., 2017; Cai et al., 2020a). It can be concluded from this, on the whole, that the use of Meitzav tests and quantitative tests does not constitute sufficient ways of evaluation in mathematics, because the achievements they measure are not tailored to the uniqueness of the students (regardless of the number of students in the classroom), do not take into account the context of the time and place in which teaching/learning takes place, nor the time of the text (relative to the time of teaching/learning). Alternative assessment could provide an appropriate response to these considerations, as it can be adapted to the learners' uniqueness, to the specific time and place where the teaching/learning takes place, and the context of the teaching/learning continuum (for example, with alternative assessment one can offer a flexible schedule, enable correcting and improving the work until achieving mastery of the material studied), because an alternative assessment allows active participation in the evaluation alongside the learning and adaptation of the study material to the unique abilities and needs of the learners (Silver & Mills, 2018). At the same time, as assessment improves and with it coping with the difficulties involved in its implementation, the quality of teaching and the essential assessment skills for improving students' learning and achievement may improve (Briggs et al., 2012; Cai et al., 2020c; Li et al., 2019; Kingston & Nash, 2011, 2012).

Among the considerations related to the ways of evaluation in mathematics, teachers focus mainly on the difficulty in revealing students' thought process and strategy that led them to the solution, as this difficulty makes it hard for them to determine how to evaluate and grade mistakes or a partial process of a solution, or a task carried out in full but with an error. These difficulties are indeed especially typical of the discipline of mathematics, and point to its complexity (Cai et al., 2018; Cai et al., 2020b). It is interesting to note that in considerations focused on the ways of evaluation themselves (not necessarily an assessment in mathematics), grading a recurring error is the consideration that is ranked lowest compared to the other considerations, as this issue is an inherent part of the evaluation in mathematics and is self-evident and it mainly characterizes quantitative tests.

An alternative assessment, which includes the possibility of interaction between the teacher and the student or all students – allows the student to present and explain the way of solution and thus overcome mistakes (choosing the way that is incorrect while is a correct solution that may be clarified in retrospect) and recurring errors (incorrect writing due to negligence or a learning disability that affects the examinee's reading/writing ability) that may impair the evaluation he receives. The difficulty in a firm and reliable assessment in mathematics, which stems from the complexity of this discipline, which includes not only unique language, procedures, and skills for solving problems solving and exercises, but also requires a deep understanding beyond the routine tasks (which can be learned by rote only) – reinforces the need to educate teachers and acquaint them with a variety of math assessment methods, understanding the considerations that must be exercised in order to choose different ways, and combine them (Schiefer et al., 2019). It is important to reiterate that evaluation using alternative assessment methods could contribute to higher achievements, improved motivation, perseverance for learning, and even to the students’ perception of efficacy (Sahin & Abali Ozturk, 2014),as well as promote the study of mathematics and its development (Chiang, 2015; Veldhuis & van den Heuvel-Panhuizen, 2020).

Conclusions

The main conclusion that emerges from the current study is that math teachers focus simultaneously on all three foci of evaluation: considerations in evaluating the teacher's specific learners, the ways of evaluation that are familiar to the teachers and available to them, and considerations that are unique to evaluation in mathematics. This conclusion is important in that, when teachers approach evaluating the math students’ knowledge, skills, and achievements, they combine in their considerations the uniqueness of mathematics assessment (such as scoring considerations) of a correct answer written with mistakes in wording or in the formal way of formulating the solution). In other words, math teachers aspire to evaluate their students correctly based on a broad picture of the students’ needs, the assessment methods available to them, and the need to focus on the evaluation difficulties that are unique to the subject of mathematics. *This conclusion* supports the approach that characterizes a reliable, dependable, and valid assessment based on data that reflects the entire learning process (Schiefer et al., 2019) combined with an alternative assessment that is essential qualitative (Chiang, 2015; Mandinach & Jackson, 2012; Sahin & Abali Ozturk, 2014; Silver & Mills, 2018; Veldhuis & van den Heuvel-Panhuizen, 2020). Also, the research findings reinforce the argument that mathematics assessment has unique considerations that are distinct and different from assessment considerations in other disciplines, and therefore mathematics teachers have a variety of assessment considerations unique to their subject. On the one hand, there is a need for specialized training for math teachers to professionalize in the uniqueness of assessment in this complex subject while they are still students of education. On the other hand, there is a need to plan and implement specialized in-service courses in assessment in mathematics as a distinct discipline for in-service teachers and create a setting for mutual collaboration with new tools for assessment in mathematics.

Learning and experiencing assessment in mathematics must be systematically integrated into the practical and theoretical education of math student teachers, together with the didactic lessons that focus on teaching and learning (the teaching-learning-assessment cycle). This assessment must include both quantitative and qualitative methods (Pellegrino, 2003; Veldhuis & van den Heuvel-Panhuizen, 2020), because complementary data that can be derived from an alternative assessment may shed light on the processes and strategies of the students' thinking and increase the teachers’ understanding of the students’ mathematical learning experiences, by providing the teachers with tools to formulate for themselves the considerations they must exercise in order to overcome their difficulties in planning and implementing an alternative assessment (Briggs et al., 2012; Cai et al., 2020c; Kingston & Nash, 2011, 2012) and their professional education in developing appropriate and reliable indicators for evaluation (Cai et al., 2018; Cai et al., 2020b; Schiefer et al., 2019). The contribution of this study is expressed in the fact that the uniqueness of mathematics assessment is emphasized – as a basis for teachers' considerations to combine varied assessment methods that will provide data that will highlight the complex aspects of learning mathematics.

References

Abali Öztürk, Y., & Şahin, Ç. (2014). The effects of alternative assessment and evaluation methods on academic achievement, persistence of learning, self-efficacy perception and attitudes. *Journal of Theory and Practice in Education (JTPE), 10*(4), 1022-1046.

Al-Nouh, N. A., Taqi, H. A., & Abdul-Kareem, M. M. (2014). EFL primary school teachers' attitudes, knowledge and skills in alternative assessment. *International Education Studies, 7*(5), 68-84. http://doi.org/10.5539/ies.v7n5p68

Beller, M. (2012). Technologies in Large-Scale Assessments: New Directions, Challenges, and Opportunities. The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research.

Beller, M. (2013). Assessment in the service of learning: Theory and practice. Ramat Gan: RAMA: The National Authority for Measurement & Evaluation in Education.

Biton, Y. & Halfon, E. (in print). Evaluation concerns of mathematic teachers and student teachers. *Practical Assessment, Research, and Evaluation.*

Black, P., & Wiliam, D. (2012). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (2nd ed., pp. 11-32). Sage. http://dx.doi.org/10.4135/9781446250808.n2

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, *31*(4), 13-17. https://doi.org/10.1111/j.1745-3992.2012.00251.x

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V. & Hiebert, J. (2018). Using Data to Understand and Improve Students' Learning: Empowering Teachers and Researchers Through Building and Using a Knowledge Base. *Journal for Research in Mathematics Education, 49*(4), 362-372. https://www.researchgate.net/publication/326219277

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V. & Hiebert, J. (2017). Clarifying the Impact of Educational Research on Learning Opportunities. *Journal for Research in Mathematics Education, 48*(3), 230-236. https://pubs.nctm.org/view/journals/jrme/48/3/article-p230.xml

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., & Hiebert, J. (2020a). Maximizing the Quality of Learning Opportunities for Every Student. *Journal for Research in Mathematics Education, 51*(1), 12-25. https://www.jstor.org/stable/10.5951/jresematheduc.2019.0005?seq=1

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., & Hiebert, J. (2020b). Working Across Contexts: Scaling Up or Replicating With Variations? *Journal for Research in Mathematics Education, 51*(3), 258-267. https://www.researchgate.net/publication/340956336

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., & Hiebert, J. (2020c). Timely and useful data to improve classroom instruction. *Journal for Research in Mathematics Education, 51*(4), 387-398. https://doi.org/10.5951/jresematheduc-2020-0056

Chiang, W. W. (2015). Ninth grade student' self-assessment in science: A Rasch analysis approach. *Procedia - Social and Behavioral Sciences*, *176*, 200-210. https://doi.org/10.1016/j.sbspro.2015.01.462

Darmody, M., Lysaght, Z., & O'Leary, M. (2020). Irish post-primary teachers' conceptions of assessment at a time of curriculum and assessment reform. *Assessment in Education: Principles, Policy & Practice, 27* (5), 501-521. https://doi.org/10.1080/0969594X.2020.1761290

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393-415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

Ediger, M. (2013). Teacher observation to evaluate mathematics achievement. *Delta-k*, *51*(1), 4-6.

Galustyan, O. V. (2017). Some methodological aspects of the evaluation of students' educational achievements at university. *International Journal of Cognitive Research in Science, Engineering and Education, 5*(1), 43-48. https://doi.org/10.5937/IJCRSEE1701043G

Kim, M. K., & Noh, S. (2010). Alternative mathematics assessment: A case study of the development of descriptive problems for elementary school in Korea. *Eurasia Journal of Mathematics, Science and Technology Education*, *6*(3), 173-186. https://doi.org/10.12973/ejmste/75238

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*(4), 28-37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

Kingston, N., & Nash, B. (2012). How many formative assessment angels can dance on the head of a meta-analytic pin: .2. *Educational Measurement: Issues and Practice*, *31*(4), 18-19. https://doi.org/10.1111/j.1745-3992.2012.00254.x

Kulm, G. (2013). Back to the future: Reclaiming effective mathematics assessment strategies. *Middle Grades Research Journal*, *8*(2), 1-10

Levy-Vered, A., & Nasser-Abu Alhija, F. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, *21*(5-6), 378-406. https://doi.org/10.1080/13803611.2015.1117980

Li, L., Fan, J., & Jin, Z. (2019). Comparing multimethod assessment of approaches to learning among preschool children: Direct measure, teacher report, and parent report. *Psychology in the Schools*, *56*(8), 1271-1286. https://doi.org/10.1002/pits.22274

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71-85. https://doi.org/10.1080/00461520.2012.667064

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.

National Council of Teachers of Mathematics. (2013). *Formative assessment: A position of the National Council of Teachers of Mathematics.* https://www.nctm.org/uploadedFiles/Standards_and_Positions/Position_Statements/Formative%20Assessment1.pdf

Nevo, D. (2002). Dialogue evaluation: Combining internal and external evaluation. In D. Nevo (Ed.), *School-based evaluation: An international perspective* (pp. 3-16). Emerald

Nevo, D. (2006). Evaluation in education. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The SAGE handbook of evaluation* (pp. 442-460). SAGE Publications Ltd. https://doi.org/10.4135/9781848608078.n20

Pellegrino, J. W. (2003). Knowing what students know. *Issues in Science and Technology*, *19*(2), 48-52.

Sahin, C., & Abali Ozturk, Y. (2014). Opinions of Prospective Teachers on Alternative Assessment-Evaluation Methods. *Kastamonu Eduaction Journal, 22*(1), 123-142.

Savickiene, I. (2011). Designing of student learning achievement evaluation. *Quality of Higher Education, 8*, 74-93.

Schiefer, J., Golle, J., Tibus, M., & Oschatz, K. (2019). Scientific reasoning in elementary school children: Assessment of the inquiry cycle. *Journal of Advanced Academics, 30*(2), 144-177. https://doi.org/10.1177/1932202X18825152

Shahbari, J. A., & Abu-Alhija, F. N. (2018). Does training in alternative assessment matter? The case of prospective and practicing mathematics teachers' attitudes toward alternative assessment and their beliefs about the nature of mathematics. *International Journal of Science and Mathematics Education, 16*(7), 1315-1335. https://doi.org/10.1007/s10763-017-9830-6

Silver, E. A., & Mills, V. L. (2018). A fresh look at formative assessment in mathematics teaching. National Council of Teachers of Mathematics.

Ültanir, E. (2012). An epistemologic glance at the constructivist approach: Constructivist learning in Dewey, Piaget, and Montessori. *International Journal of Instruction, 5*(2), 195-212

Veldhuis, M., & van den Heuvel-Panhuizen, M. (2014). Primary school teachers' assessment profiles in mathematics education. *PLoS ONE, 9*(1), Article e86817. https://doi.org/10.1371/journal.pone.0086817

Veldhuis, M., & van den Heuvel-Panhuizen, M. (2020). Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Mathematics Education Research Journal, 32,* 449-471. https://doi.org/10.1007/s13394-019-00270-5

Watt, H. M. G. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in Sydney, Australia. *Educational Studies in Mathematics, 58*, 21-44. https://doi.org/10.1007/s10649-005-3228-z

Zhao, X., van den Heuvel-Panhuizen, M., & Veldhuis, M. (2018). Chinese primary school mathematics teachers' assessment profiles: Findings from a large-scale questionnaire survey. *International Journal of Science & Mathematics Education*, *16*(7), 1387-1407. https://doi.org/10.1007/s10763-017-9841-3

Author Information

Ester Halfon

https://orcid.org/0000-0002-7717-352X

Shaanan College of Education

Haifa 2640007

Israel

Contact e-mail: *halfonester@gmail.com*

Yaniv Biton

https://orcid.org/0000-0003-2329-8877

Shaanan College of Education

Haifa 2640007

Israel