Exploring metrics: elementary mathematics teachers' evaluation of digital geometry assessment activities

Hassan Ayoob[1] and Shai Olsher[2]

University of Haifa, Faculty of Education, Israel; [1]hassan.ayoob@gmail.com; [2]olshers@edu.haifa.ac.il

This study concentrates on metrics that guide teachers in perceiving Digital Formative Assessment (DFA) as suitable for formative assessment for students and aligning with the curriculum. This is part of a larger study that explores the efficiency of a digital formative assessment platform for elementary school geometry. The research utilized open coding to analyze responses from a teacher questionnaire, exploring elementary math teachers' perspectives on DFA. This research was conducted with nine mathematics teachers from diverse Israeli elementary schools. Using a teacher questionnaire and 12 DFA activities, the study identified 11 codes categorized into three categories: Information provided, type of task, and student interaction.

Keywords: Digital formative assessment, elementary math teachers' perspectives, curriculum alignment.

Introduction and theoretical background

Formative assessment is a continuous, multifaceted process integrated into the daily dynamics of teaching and learning. It unfolds through ongoing interactions between teachers and students, wherein teachers adjust their instructional methods and activities based on assessment information (Black & William, 2009). The primary goal is to enhance the learning processes and improve student outcomes. According to Black et al. (2004), the essence of formative assessment lies not only in the assessments themselves, but also in the roles they play in supporting student learning and providing evidence for adapting teaching methods to address specific learning needs. Approaching formative assessment from this functional perspective underscores that its successful implementation hinges on the learning approach adopted and the adept use of knowledge, skills, and strategies by teachers in executing intricate pedagogical processes (Webb & Jones, 2009). For effective formative assessment, teachers' proficiency in regularly collecting and interpreting student learning data is crucial. This complex task extends beyond simple data collection, incorporating timely and constructive feedback, understanding of student learning goals, and tailoring teaching methods accordingly (Black & William, 1998; Heritage, 2007). Central to formative assessment is the role of feedback, which should be continuous, precise, and focused to guide both educators and learners. Feedback is vital for reflective learning, helping students recognize their progress and areas for improvement. Feedback also provides teachers with valuable insights about their teaching methods, by identifying areas where students struggle and understanding students' learning styles. This process helps teachers refine their instructional strategies and make informed decisions to enhance their teaching (Black & William, 1998; Hattie & Timperley, 2007; Heritage, 2007). While many researchers and teachers are aware of the importance of formative assessment for improving the quality of learning and increasing achievement, several factors, such as lack of time for implementation and challenges in adapting tasks, negatively influence formative assessment’s implementation in the classroom (Brown, 2003).

In addition, many researchers agree that teachers' perceptions are one of the main factors impacting implementation in the classroom (Brown, 2003; Black & William, 2009). Assessment, especially formative assessment, is important in the mathematics teaching and learning process, as it helps realize the curriculum's objectives by adapting teaching (Black et al., 2003; Millar, 2016). In recent years, teaching with digital environments in general and geometry in particular has increasingly relied on technology. By offering dynamic forms of mathematical representation of different concepts, dynamic geometry environments (DGEs) provide students with access to mathematical concepts that they have not previously perceived (Leung, 2008; Butler et al., 2010). The suggested DFA activities are designed to effectively support the identification, classification, and analysis of student work methods to the relevant stakeholders (Ayoob & Olsher, 2023). The activities include rich tasks with an infinite number of correct solutions, in the form of the example-eliciting task (EET) (Olsher et al., 2016). These (DFA) tasks feature interactive feedback that highlights the students' exploration and reasoning beyond right or wrong answers. Such design promotes an in-depth interaction with geometric concepts, enabling students to articulate and reflect on their thought processes (Olsher et al., 2016), thereby facilitating a critical examination of their reasoning strategies (Stacey & William, 2013). This type of task involves student-centered assessment that can provide teachers and students with the characteristics of each student’s work (Olsher, 2022).

Methodology

The qualitative methodology in this research involves using open coding to analyze responses from an open-ended teacher questionnaire. The focus of this study is to systematically examine elementary mathematics teachers’ perspectives regarding computerized activities, specifically digital formative assessment (DFA).

This study is part of a larger research project that explores the efficiency of DFAs for elementary school geometry. Our goal in this report is to describe the metrics according to which elementary mathematics teachers perceive digital formative assessment activities as appropriate for their students and the content of the curriculum. To achieve this goal, we seek to answer the following research questions: What metrics do math teachers use when evaluating the suitability of a digital formative assessment (DFA) to the content and skills detailed in the curriculum?

Research setting

In this study, we analyzed the responses gathered through an open-ended questionnaire. Using an open coding qualitative method, we systematically examined teachers' insights to identify and categorize the metrics they provide. We aimed to address our research question by delineating emerging patterns and themes. This approach enabled us to uncover rich, context-specific information about teachers' perspectives on the effectiveness and alignment of DFA's with curriculum goals and student skills.

Population

Nine mathematics teachers from five different Arabic-speaking elementary schools in northern Israel. All schools teach according to Israel’s national curriculum. The participants are a diverse group of teachers with varying educational backgrounds and levels of experience. The majority held master's degrees, with teaching experience ranging from 8 to 25 years, in the 3rd to 6th grades. Collectively, the teachers implemented 69 digital formative assessment activities (and answered the accompanying questionnaires), showing active participation in the research.

Research tools

To examine the metrics according to which elementary mathematics teachers evaluate DFA activities as appropriate for students and the curriculum content, we used a teacher questionnaire and DFA units designed in the STEP-MFA system.

The **12 activities** were designed according to the design principles set out in previous research (Ayoob & Olsher, 2023). The activities align with the textbooks and with the recommended instruction and distribution of the topic to teaching hours according to the curriculum. The tasks were designed in the STEP-MFA environment, and the automatically assessed characteristics of student submissions were designed according to the misconceptions and perceptions in previous literature and the common errors for each subject.

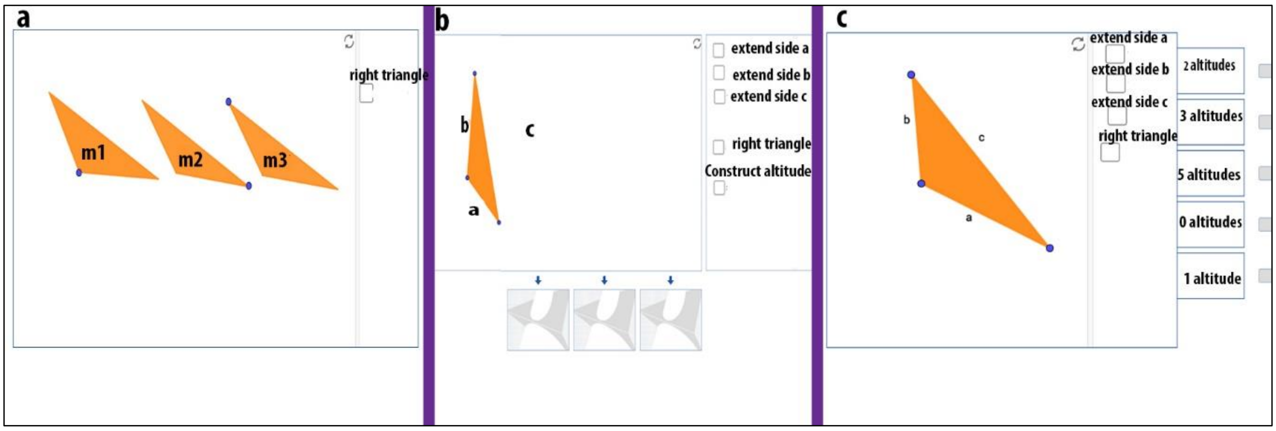

Figure 1: Three tasks of the obtuse triangle altitude activity: (a) construct altitudes from the marked vertex, (b) construct three different examples of obtuse triangles and their altitudes, and (c) determine how many altitudes are outside the triangle?

The following are examples of DFA activity: The mathematical topics studied in this study were the triangle's altitude and polygon area. The activity about the altitudes of an obtuse triangle is shown in Figure 1. The first task (figure 1a) shows three instances of the same obtuse triangle. In each instance, the student is asked to drag the bold vertex to construct an altitude from that vertex and submit it. The triangles are static, and students can drag only the bold vertices. In the second task (Figure 1b), students were asked to construct three different examples of obtuse triangles and then construct one altitude by dragging a point from a vertex. Students can drag the vertices of a triangle to create obtuse triangles of their choices. In the third task (Figure 1c), students are asked to specify how many altitudes can be constructed that are external to the static obtuse triangle and to construct them again by dragging points from the vertices to “stretch” segments. On the left side of each task, students have the option to select and use several tools, such as the option to display the lines extending the sides or a right-angled triangle ruler.

The **questionnaire** consisted of 12 questions, each featuring a dual structure. The first part involves a query with responses based on a linear or multiple-choice format. The subsequent section required the participants to explain their selection. For this study's purpose, we focused on six open-ended questions pertaining to qualitative analysis, concentrating solely on verbal explanations. Question 3 directly addressed indices, inquiring about the metrics used to determine the suitability of a task as a formative assessment activity. Additionally, Question 6 sought an explanation for choosing the tasks deemed most suitable for the students. Question 7 sought an explanation of the rationale behind selecting tasks that were considered unsuitable for students. Question 8 sought an explanation for choosing tasks that provided information about students. Question 9 aimed to understand the reasoning behind choosing tasks that did not offer information about the students. Finally, Question 12 explored the preference for a specific report among the four reports available on the platform.

Research setting

Preparatory stages were undertaken before introducing activities to the classroom for both teachers and students on the STEP platform. Throughout the academic year, our focus was on ensuring that teachers promptly conducted assessments after teaching each sub-topic. The system diligently recorded all the submissions for every student. After the assessment and after any discussions with students about their submissions, the teachers completed the questionnaire. In this questionnaire, they elaborated on their perspective regarding the tasks as formative assessment tasks and explained why they deemed them appropriate.

Following the completion of all activities and the collection of teachers' responses, we initiated a coding process for teachers' answers, categorizing them based on their characteristics. The coding process and questionnaire were validated by the Mathematics Education Research and Innovation Center (MERI) team comprising graduate students, curriculum developers, and researchers in mathematics education.

Data collection and analysis

The data for this study were obtained from the responses of nine teachers to six open-ended questions in the questionnaire, following the completion of each of the 12 research activities.

The data analysis process in this study followed a qualitative approach, primarily utilizing open coding to extract meaningful insights from the responses obtained through a teacher questionnaire. Initially, teachers' answers were transcribed and organized into sheets by question. Open coding involves systematically examining and labeling data segments with descriptive codes, capturing the essence of each unit. Through constant comparison, new data were compared with existing codes to refine the categories and identify patterns and variations. Categories were developed to represent common themes emerging from teachers' insights, progressing from specific codes to more generalized concepts. To ensure the validity of our analysis, a validation process was implemented, cross-checking preliminary findings with the MERI team and validating over 15% of the data. In all cases, disagreement discussions led to consensus and, in some cases, modification of the coding scheme.

Results

In our exploration of teachers' responses to the questionnaire, a comprehensive dataset of 69 statements for each question was collected. The coding process yielded 11 distinct codes. Subsequently, these codes were aggregated and structured into three overarching categories, as detailed in Table 1. **Information provided**, **type of task**, and **student interaction.** A detailed breakdown of these categories is provided below.

Table 1: Categories and codes for Teachers' responses on the suitability of activities as formative assessment

category	codes
Information provided	Classroom or group information
	Information about the student (Individual)
	Correctness
	Student work methods and misconceptions
Type of task	Curriculum objectives
	Difficulty levels
	Variety of tasks
	Suitability for a diversity of students
Student interaction	Task modification
	Number of submitted examples
	Thought development

Following table 1, and to elucidate the categories and their corresponding sub-categories, we will outline each one followed by its sub-categories. For each sub-category, we'll provide an example of a teacher's verbal response. This approach aims to clarify the rationale behind our categorization and selection process, offering insights into the specific teacher feedback that influenced our decisions.

Information provided

This category captures the depth and variety of information that DFA activities offer teachers about student learning and comprehension. Teachers looked at the information provided by the activity for insights about their class or a group of students. "In my view, if I need to teach the topic again because of students' mistakes, it suggests that the tasks are fitting for assessment activity" (Teacher A, answer to Question 3). Capturing nuanced details specific to individual students. "The activity gives me an in-depth idea of the student's level, especially after completing all the tasks in the activity" (Teacher N, Q 3).

This activity allows teachers to obtain information on the correctness or errors in students' responses. "As a teacher, I will have the knowledge of who has answered correctly and who has not, and I will also be aware of how often certain answers are given. This will enable me to effectively deal with the outcomes" (Teacher B, Q 12). Analysis of the varied methods employed by students in their work methods. "It offers the opportunity to understand all the misconceptions" (Teacher L, Q 8).

Type of task

Teachers assess the relevance of DFA activities based on their alignment with curriculum goals, their challenge level, and the variety they offer. The activity matches and checks the curriculum objectives in every aspect. "The activity effectively checks a student's understanding of whether two segments are perpendicular, not only the segments themselves but also their extensions. It assesses comprehension across various levels: basic understanding, knowledge, identification skills, and application proficiency" (Teacher L, Q 3). Perceived difficulty levels of the assigned tasks. "The activity is suited as it caters to both advanced and less proficient students" (Teacher L, Q 3). Diversity and range of tasks. "The activity includes a variety of tasks designed to cater to and challenge the different skill sets of the students" (Teacher D, Q 3). Appropriateness of tasks for specific student groups. "In this task, students are required to provide three examples. However, from my experience, asking for ten examples is challenging. It is probable that only a few students would be able to present ten examples, but those who do show an ability to generalize" (Teacher D, Q 8).

Student interaction

This category reflects how DFA activities facilitate student engagement through interactive and thoughtful task design. Tasks that enable students to modify the given mathematical context while constructing an example. "The second task allows the student also to build and vary the types of triangles he chooses" (Teacher A, Q 3). Preferences and considerations surrounding the number of prompted examples within a given task. "When a student submits numerous examples, their response becomes clearer and more comprehensible" (Teacher R, Q 8). Tasks for the potential to foster the development of critical thinking skills. "Different examples develop students thinking, so we will have an assessment tool that also enables learning" (Teacher R, Q 6).

Summary and discussion

The categorization of teachers' responses into three primary categories provides an understanding among mathematics teachers of the multifaceted considerations involved in effectively leveraging formative assessment tools, reflecting a deep understanding of the multifaceted considerations required to effectively leverage these tools in education.

Exploring the **information provided** by the assessment task category, teachers acknowledged the importance of designing formative assessments for all students in the classroom or student groups, recognizing the diverse learning needs inherent in different educational stages. The focus on individual student characteristics, especially technological proficiency, underscores commitment to personalized approaches in formative assessment. Additionally, the acknowledgment of addressing and correcting misconceptions highlights the role of formative assessment in fostering accurate understanding and facilitating future teaching and learning.

Within the **type of task** category, teachers articulate their perspectives on the design and structure of formative assessment tasks. Emphasis on task structure indicates an awareness of how instructional design influences student engagement. Awareness of developing students' thinking skills reflects an approach that extends beyond measurement to actively cultivate critical thinking abilities. Furthermore, the acknowledgment of the challenge in balancing difficulty levels underscores the commitment to providing tasks that appropriately challenge students while ensuring accessibility. The incorporation of a variety of tasks and the consideration of task suitability for specific student populations reveals the aim of creating diverse and inclusive learning experiences.

The **student interaction** category shows that teachers recognize the value of providing flexibility in both the number of examples and the mathematical context that can be altered to fit the interest or level of difficulty preferred by the student. Consideration of the number of examples indicates a thoughtful approach to balancing the assessment's depth and breadth.

These insights reflect an understanding of formative assessment in line with the principles discussed by Black and William (1998) and Heritage (2007).

In general, these categories contribute valuable insights into the ongoing discourse on best practices in formative assessment, particularly in the context of digital tools and technology integration. These insights reveal a holistic perspective on formative assessment practices. Some of the teachers also showed awareness of the importance of receiving information about the student beyond his mistakes, but also of the work method, which is aligned with the capacities of the STEP platform (Olsher et al., 2016).

In addition to the identified categories, it is important to explore teachers' attitudes towards students' automatic feedback and their integration into formative assessments. Teachers did not refer to this type of feedback in the form of personal reports. This observation is significant considering the established role of timely feedback in formative assessment (Hattie & Timperley, 2007). The absence of a focus on teachers' responses may indicate an area for further exploration and emphasis on teacher training. Additional aspects that did not receive attention in this study included the use of technology to facilitate collaborative work among students, including pairings and peer learning experiences. Understanding how teachers harness digital tools to foster collaborative learning environments and whether they endorse peer-driven feedback processes can offer additional insight into digitally enhanced formative assessment practices. Furthermore, the observation that some teachers showed a tendency to focus primarily on identifying where students went wrong underscores a critical area for development in teacher training. Recognizing and addressing the underlying thought processes behind both correct and incorrect student responses can lead to more nuanced teaching strategies that better support student learning.

References

Ayoob, H., & Olsher, S. (2023, July). Curriculum-aligned digital formative assessment for elementary school geometry. In Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13) (No. 4). Alfréd Rényi Institute of Mathematics; ERME.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Black, P., Harrison, C., Marshall, L. B. & William, D. (2004). Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8–21. https://doi.org/10.1177/003172170408600105

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability (formerly: Journal of personnel evaluation in education), 21, 5-31. https://doi-org.ezproxy.haifa.ac.il/10.1007/s11092-008-9068-5

Brown, G. T. (2003, November). *Teachers' instructional conceptions: Assessment's relationship to learning, teaching, curriculum, and teacher efficacy*. At a joint conference of the Australian and New Zealand Associations for Research in Education (AARE/NZARE), Auckland (Vol. 28).

Butler, D., Jackiw, N., Laborde, J. M., Lagrange, J. B., & Yerushalmy, M. (2010). Design for transformative practices. In C. Hoyles & J.-B. Lagrange (Eds.), *Mathematics Education and Technology-Rethinking the Terrain* (pp. 425–437). Springer.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112. https://doi-org.ezproxy.haifa.ac.il/10.3102/003465430298487

Heritage, M. (2007). Formative assessment: What do teachers need to know and do?. *Phi Delta Kappan*, *89*(2), 140-145. https://doi-org.ezproxy.haifa.ac.il/10.1177/003172170708900210

Millar, R. (2017). Using assessment materials to stimulate improvements in teaching and learning. In K. Hahl, K. Juuti, J. Lampiselka, A. Uitto, & J. Lavonen (Eds.), *Cognitive and Affective Aspects in Science Education Research* (pp. 31-40). Springer. https://doi.org/10.1007/978-3-319-58685-4_3

Leung, A. (2008). Dragging in a dynamic geometry environment through the lens of variation. *International Journal of Computers for Mathematical Learning*, *13*, 135-157.

Olsher, S., Yerushalmy, M., & Chazan, D. (2016). How might the use of technology in formative assessment support changes in mathematics teaching? *For the Learning of Mathematics*, *36*(3), 11–18. http://www.jstor.org/stable/44382716

Olsher, S. (2022). Te(a)Ching to collaborate: Automatic assessment-based grouping recommendations and implications for teaching. *Proceedings of the 15th International Conference on Technology in Mathematics Teaching (ICTMT 15)*, (pp. 214–223). Danish School of Education, Aarhus University. https://doi.org/10.7146/aul.452

Stacey, K., & Wiliam, D. (2013). Technology and assessment in mathematics. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick, F. Leung (Eds.) *Third International Handbook of Mathematics Education* (Vol.27, pp. 721–751). Springer. https://doi.org/10.1007/978-1-4614-4684-2_23

Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. Assessment in Education: *Principles, Policy & Practice, 16(2),* 165-184. https://doi.org/10.1080/09695940903075925