MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(11), em2526
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/15482

Investigating students' arguments with real-life functional situations throughout a sequence of collaborative activities

Michal Ayalon [1*], Karina J. Wilkie [2], Rina Swaid [1]

[1] University of Haifa, Haifa, ISRAEL
[2] Monash University, Melbourne, VIC, AUSTRALIA

Received 28 May 2024 ▪ Accepted 19 September 2024

Abstract

Recent years have seen increasing interest in providing students with opportunities for developing important argumentation skills in the mathematics classroom. Social interactions with peers to critique alternative ideas, justify arguments, and build consensus, have been found to promote deep thinking and meaningful development of concepts. In this study we explored 9th-grade students' interactions during a sequence of specifically designed argumentation tasks on real-life functional situations to investigate the appropriateness of their arguments. Data were collected from the students' written task responses, student reflections, small-group observations, individual interviews with the group members, and teacher interviews. Analysis of the level of appropriateness of the students' individual and group written responses in each activity focused on three aspects: identifying variables, forming relations between them, and noticing contextual features of the real-life situation. We found evidence of students grappling with selecting two suitable variables and with conceptualizing the nature of their relation. It also appeared that aspects of the students' social interactions played a role in students ignoring correct arguments and accepting incorrect arguments. We discuss implications for small-group argumentation and suggest avenues for future research.

Keywords: argumentation, productivity in argumentation, functions, graphing real-life situations, secondary mathematics

INTRODUCTION

Argumentation is one of the main processes associated with the competencies that can help students to deal better with the challenging demands of 21st century life (e.g., Toh & Kaur, 2016). Such competencies include critical thinking, communication and collaborative skills, and characteristics such as ownership of one's learning, curiosity, courage and flexibility. Learning to voice arguments, exchange ideas, listen attentively to others' ideas, and critically evaluate the strengths and weaknesses of different perspectives–all argumentation activities–can play an essential role in developing students' competencies (Chua, 2016). Research in mathematics education suggests that participation in argumentation activities that require the student to explore, confront, and evaluate alternative positions, voice support or objections, and justify different ideas and hypotheses, promotes meaningful understanding and deep thinking (Asterhan & Schwarz, 2016; Francisco & Maher, 2005; Weber et al., 2008; Wood et al., 2006). This view is reflected in recent educational reform documents all over the world and in Israel, in particular, that highlight argumentation as one of their important goals for students (e.g., Common Core State Standards Initiative [CCSSI], 2010; Ministry of Education, 2020, 2024).

Studies have highlighted that argumentative activity in mathematics is demanding and requires teachers and students, alike, to possess specific intellectual and social skills (Ayalon & Even, 2016; Staples, 2014). This is often emotionally challenging and poses the difficulty of engaging students in productive argumentation that is not teacher-centered (Mueller et al., 2014). There is a need to understand more about students' and teachers' experiences and perspectives (Chazan, 1993; Corbett & Wilson, 1995) to operate wider-scale implementation of argumentation in mathematics classrooms. This study is

✉ mayalon@edu.haifa.ac.il **(*Correspondence)** ✉ karina.wilkie@monash.edu ✉ rhayek@campus.haifa.ac.il

Contribution to the literature

- During individual and small-group activities on graphing real-life functions students were found to identify differing possible variables and make sense of how they related to each other using contextual features of the situation, with varying levels of mathematical appropriateness.
- Various aspects of the students' social interactions with each other were found to influence students' inattention to correct arguments and acceptance of incorrect arguments.
- It is important for teachers to manage both the level of mathematical complexity of collaborative tasks as well as monitor the social dynamics involved in peer critiquing to support productive argumentation.

part of a broader project that aims to investigate students' productivity and engagement in mathematical argumentation in relation to the implementation of purposefully designed, student-centered tasks. It incorporates some novel features for mathematics learning that require fine-grained research. These include argumentation tasks where students collaboratively evaluate exemplars, co-construct quality criteria for assessing arguments, and self-assess (Burkhardt & Swan, 2012)–activities conducive for learning and facilitative of formative assessment that are not widely researched in secondary mathematics contexts and remain under-theorized (Schneider & Randel, 2010). A series of tasks and associated templates for structuring students' work have been written for this pilot project (see **Appendix A** for an example). This in-depth qualitative study was conducted in the context of a constructive teaching experiment with a class of secondary students.

In this study, we concentrate on students' *productivity* in mathematical argumentation (Schwarz & Baker, 2017). Productivity encompasses two dimensions: *structural* and *dialogic* (McNeill & Pimentel, 2010). Structurally, productivity in argumentation pertains to the appropriateness of arguments, recognizing that what is considered sophisticated may vary depending on the specific classroom or context of the lesson (Yackel & Cobb, 1996). Dialogically, it involves students presenting diverse perspectives, critically and respectfully engaging with others' ideas, striving for consensus, and justifying their reasoning. Although productivity is considered important in argumentation, in the literature there seems to be very little research on students' productivity in mathematical argumentation. This study addresses this lacuna, with a particular focus on the structural component of productivity. It examines the appropriateness of the students' arguments through participation in a sequence of collaborative argumentation activities. The mathematical context for the argumentation task sequence in this study is drawing graphs according to verbal descriptions of non-linear real-life functional situations, followed by evaluating the match between some provided graphs and each situation. The tasks focus on functions concepts: identifying the variables, building the relationship between them (covariation), and noticing contextual features from each real-life situation. The students' familiarity with everyday events in the real world can serve as a basis for learning how to interpret a graph of a function (Goldenberg, 1987) but graphs compatible with real-life situations seem to present students with particular difficulties (Ayalon et al., 2016b).

THEORETICAL BACKGROUND

This section focuses on two main issues:

(1) theoretical perspectives on argumentation with a specific focus on productivity in mathematical argumentation, and

(2) students' learning of functions with real-life situations.

Theoretical Perspectives on Argumentation

A commonly accepted definition of argumentation in educational research is that of van Eemeren and Grooendorst (2004) in which they assert that argumentation is "a verbal, social, and rational activity aimed at convincing a reasonable critique of the acceptability of a standpoint by putting forward a constellation of propositions justifying or refuting the proposition expressed in the standpoint" (p. 1). It involves building claims, providing evidence to support the claims, and evaluating evidence to judge the validity of the claims. If incorporated as part of classroom discourse, it affords a venue for the articulation of alternative ideas, reflection, and reasoning (Chin & Osborne, 2010). Descriptions in the literature of argumentation that is 'fruitful' for learning refer to "balances between critical reasoning and collaborative knowledge construction" (Asterhan & Schwartz, 2016, p. 167). This type of argumentation–termed *deliberative argumentation* (Felton et al., 2009)–is characterized by critically and respectfully listening to others' ideas, searching for alternative ideas, reaching consensus, and accountability to reasoning.

Productivity in mathematical argumentation

An argumentation that is productive for students learning to argue and arguing to learn has two important meanings: *structural* meaning and *dialogic* meaning (McNeill & Pimentel, 2010). The *structural* meaning of argumentation focuses on the aspect of discourse in which a claim is supported by an appropriate justification. We adopted parts of Toulmin's (1958) model of argumentation for this structural definition (Reuter, 2023; Zambak & Magiera, 2020). We define the structure as consisting of three essential components: *claim*, *data*, and *warrant*. The *claim* (C) is the conclusion that answers the question or problem. The *data* (D) are the foundations on which the argument is based, the relevant evidence for the claim. The *warrant* (W) justifies the connection between data and conclusion by, for example, appealing to a rule or a definition. Productivity in argumentation in structural terms refers to the appropriateness of arguments, keeping in mind that what is deemed sophisticated might vary according to the specific classroom or context being discussed in the lesson (Yackel & Cobb, 1996). The *dialogic* meaning focuses on the interactions between individuals when they attempt to generate and critique each other's ideas (McNeill & Pimentel, 2010). This meaning aligns with *deliberative* argumentation and with the capabilities described in the aforementioned curriculum and policy documents. Productivity in argumentation in dialogical terms refers to students' raising different views and ideas, critically and respectfully listening to other's ideas, seeking consensus, and accounting for their reasoning. Although productivity is considered central to argumentation (Schwarz & Baker, 2017), there seems to be scarce research in the literature on students' productivity in mathematical argumentation. This study addresses this lacuna, with a particular focus on the structural component of productivity. It examines the appropriateness of the students' arguments through participation in a sequence of collaborative argumentation activities in the mathematical context of constructing and critiquing graphs of real-life functional situations. We focus on three aspects of appropriateness of students' arguments: identifying variables, forming relations between them, and noticing contextual features (Ayalon et al., 2016a, 2016b).

Students' Learning of Functions with Real-Life Situations

The mathematical context for the argumentation tasks in this study is drawing graphs according to verbal descriptions of non-linear situations, followed by evaluating the match between the given graphs and the situations. The situations focus on identifying the variables, forming the relationship between them (in a particular covariation) and noticing contextual features. The notion of a variable is fundamental to understanding functional relationships and graphical representations and is a prerequisite for making sense of covarying quantities. The construction of axes is not straightforward (Leinhardt et al., 1990), particularly when the variables are unusual for the learner. If these axes are provided, students do not need to identify variables and decide how to represent them. Once the variables in a situation are understood, students will need to understand variability, the way dependent and independent variables change together (e.g., Slavit, 1997). The students' familiarity with everyday events in the real world can serve as a basis for learning how to interpret a graph of a function (Goldenberg, 1987) but graphs compatible with real-life situations seem to present students with particular difficulties.

The most prominent difficulty is interpreting a graph as a literal picture of a situation (e.g., Clement, 1985; Janvier, 1981). Students deal more effectively with graphs of functions when one variable depends on time (Janvier, 1981; Thompson, 1994). The familiarity of time, plus its unidirectional nature (time only increases), seem to account for this. Students need only to comprehend how one variable (the non-time variable) varies; the time variable is implicitly taken for granted. This can lead students to ignore the meaning of the independent variable or to interpret the shape of a graph of a situation as a picture of that situation (e.g., Leinhardt et al., 1990; Schultz et al., 1986). Other challenges may be focusing on one variable only, choosing one relevant variable and one irrelevant variable and forming an irrelevant relation, choosing relevant variables but forming an inadequate relation between them, and not taking all contextual features into account (Ayalon et al., 2016a, 2018). Complex variables, such as speed, density, and price per unit, are also known to pose significant challenges (Ayalon et al., 2016a).

RESEARCH CONTEXT AND TASK SEQUENCE DESIGN

Task design for argumentation refers to the design decisions educators make when a learning task is developed, including the choice of topic for argumentation, the way in which the topic is presented, the types of resources that learners will be able to access, the group formation, and the sequencing of activities in which learners will participate (Schwarz & Baker, 2017). Design decisions regarding these particulars affect the likelihood that students will engage in argumentation (e.g., Andriessen & Schwarz, 2009; Jimenez-Aleixandre, 2007). For example, the content of an argumentation task should be problematic (Engle & Conant, 2002). This can be achieved by using problems that do not have clear solutions agreed upon by all experts. Another way to stimulate argumentative discourse is to encourage questions, suggestions, challenges, and other intellectual contributions, instead of expecting students simply to assimilate facts and procedures (Lemke, 1990). The research advocates for students to work in small groups, to discuss with peers, to make arguments, to criticize others' arguments, and to draw conclusions (Mueller et al., 2014).

A common approach in science education for fostering argumentation involves students working in small groups while considering argumentation as a shared object–a schema–to be built and discussed together (McNeill & Krajcik, 2007; Schwarz & Baker, 2017). The argument schema, such as that proposed by Toulmin (1958), becomes a tool for clarifying student reasoning in terms of a set of categories, or a meta-language of the domain, such as 'data' or 'evidence', 'claim', and 'warrant' (Toulmin, 1958). The schema provides not only a means for structuring students' discussions, but also, and particularly, a means for structuring the task in which the students are engaged. Thus, students collaboratively work on a task that requires them to represent their reasoning and the problem solution in the form of a Toulminian argument schema. It can enable their discussions to be structured argumentatively with respect to the schema (e.g., Chin & Osborne, 2010). Providing a rationale for creating arguments and providing students with criteria for the construction and evaluation of arguments, has been shown to improve argumentation skills as well (e.g., Erduran & Jiménez-Aleixandre, 2008). Given the promising findings in recent literature about the improvements in students' learning in the science classroom (e.g., Cavagnetto & Kurtz, 2016), we believe that the approaches developed for the science classroom have the potential to be investigated in mathematics classrooms as well.

Based on this information, the use of matching verbal situations and graphs, the Toulminian argument schema, example-graphs assessment, and collaborative argumentation activities that enable building criteria for the construction and evaluation of arguments is expected to encourage students' productivity in mathematical argumentation. This study will examine the appropriateness of the students' arguments through participation in a sequence of collaborative argumentation activities in the mathematical context of constructing and critiquing graphs of real-life functional situations. In particular, the study will address the following research question:

What characterizes the level of appropriateness in students' arguments throughout participation in a sequence of collaborative mathematics argumentation activities on graphing real-life situations?

A series of mathematical tasks and associated templates were developed for the study. These include argumentation tasks where students collaboratively construct arguments, evaluate example-graphs, and learn to address appropriate criteria for the evaluation of arguments. The activities were employed in eight lessons.

Figure 1. Container-1 (Ayalon et al., 2021)

The Learning Task Sequence

The main part of the activity sequence featured three cycles of two lessons, each cycle focusing on a different real-life situation.

- Lesson #1. Individual pre-sequence assessment.
- Lessons #2-3–Cycle 1. "Watering a plant" situation: Jack forgot to water his peas. The seedling dried up and its growth decelerated for a while. Then Jack remembered to water the seedling, and its growth accelerated.
- Lessons #4-5–Cycle 2. "Filling a container" situation: A tap is turned on hard, and water rushes into the container in **Figure 1** at a constant rate.
- Lessons #6-7–Cycle 3. "Pricing cakes" situation: Karin, a cake shop owner, must decide on the price at which to sell a new chocolate cake to customers. If the price is too low, she will lose money, but if the price is too high, fewer chocolate cakes will be sold, and she will lose money. She needs to choose a reasonable price so that she can sell enough cakes to make a profit.
- Lesson #8. Individual post-sequence assessment.
- Post-sequence interviews (students and the teacher) (see **Appendix B** for the interview schedules and **Appendix A** for an example for argumentation tasks).

In this paper we draw on data from the sequence of three cycles (lessons #2-7) and student and teacher post-sequence interviews. **Table 1** summarizes the structure and components of these three cycles.

The Toulminian schema provided a means for structuring the students' argumentation. In their written products, the students were asked to justify their drawn graph (the claim), by specifying the data (*a specific piece of evidence from the story of the situation*) and the Warrant (a *mathematical reason why my graph matches the piece of*

Table 1. Structure of each two-lesson cycle (lessons #2 & 3, 4 & 5, and 6 & 7)

Lessons and activities
Lesson i: individual graph construction, small group comparison, group construction
Activity 1. Individual students construct own graph of the situation.
Activity 2. In small groups, students compare their graphs and write down differences; students discuss to reach consensus and then draw a final group graph with a list of data and warrants.

Table 1 (continued). Structure of each two-lesson cycle (lessons #2 & 3, 4 & 5, and 6 & 7)

Lessons and activities
Lesson ii: critique of 3 sample graphs, small group revision, whole class comparison
Activity 3. In same group as lesson i, students critique 3 provided graphs on handout listing strengths and weaknesses.
Activity 4. Students re-visit their group graph with a list of data and warrants from last lesson and make decision to keep or revise; group writes justification for decision.
Activity 5. Final graphs of each group are displayed on board (students draw on whiteboard or stick up A3 pages); representatives of each group present the graphs to the whole class.
Activity 6. Whole class compares similarities and differences of groups' graphs.
Activity 7. Individuals reflect on their experiences.
Post-cycle activity: Teacher and focus group students participate in individual post-cycle interviews about some aspect of a group's discussion.

evidence). See **Appendix A**, activity 2, for the handouts which are structured with this Toulminian scheme.

All situations focused on identifying variables, forming relationships between them, and noticing contextual features. We chose the sequence of situations with increasing level of difficulty: from more familiar variables (Janvier, 1981; e.g., time, height) to less familiar and compound variables (Leinhardt et al., 1990; e.g., profit and growth), and from increasing to parabolic (increasing and decreasing) shapes of graphs.

METHODOLOGY

This study employed an in-depth qualitative case study (Creswell, 2007) as part of an overall educational design research program on secondary students' argumentation in mathematics. In the following three sub-sections we provide information on the participants, data collection and the analysis process.

Participants

The study took place in a 9th grade class from a non-selective government school in northern Israel. The whole class participated in the sequence of argumentation tasks, but our study described in this article focused on a small group of six students (Anna, Eva, Roni, Liam, Omer, and Tom, pseudonyms). The size and membership of the group was chosen by the class teacher in terms of their talkativeness and the likelihood of some students being absent for school rehearsals. The study took place in a school with middle-to-high levels in mathematics achievement, to minimize the risk of students finding the tasks too difficult, while at the same time wanting them to be sufficiently unfamiliar with the tasks to stimulate deliberative argumentation. The mathematics teacher was informed of the research process and how the classroom activities and particularly her role in facilitating the class discussion at the end of each cycle.

Data Collection Methods

The goal of the study was to analyze productivity in argumentation from a structural perspective, meaning, the increase in sophistication of arguments over time (Schwarz & Baker, 2017). To this end, we focused on examining the level of appropriateness of the students' arguments. Multiple data sources were used to triangulate students' experiences, teacher perceptions, and researcher observations (Creswell, 2007). The main data source for investigating the level of appropriateness of the students' arguments was the students' written task responses. These included individual graphs and explanations, the group's "best graph" and explanations, group evaluation of given example-graphs, and the group's revised and updated graph and explanations.

Other data sources included transcripts of the videoed group interactions and individual student interviews at the end of each cycle. The interviews focused on the students' experiences with the tasks and their perspectives on structural aspects of productivity in argumentation, and any 'critical' events noted by the researcher during the observation, related to, e.g., making claims, justifying, challenging, and disagreeing. In addition, we conducted interviews with the teacher at the end of each cycle and a final teacher interview. The aim was to explore the teacher's perspective on what she noticed about the class's response to the tasks, any surprising/ interesting/ difficult moments, the learning of the focus group, and her perception of the students' progress over the lesson sequence. We analyzed data on the students' discussions about variables, relations between variables, and contextual features to seek confirmation of our coding of the students' written products (Creswell, 2007). The goal was to find more evidence by triangulating data sources to seek confirmation of the classifications of the students' written products (Creswell, 2007).

Data Analysis

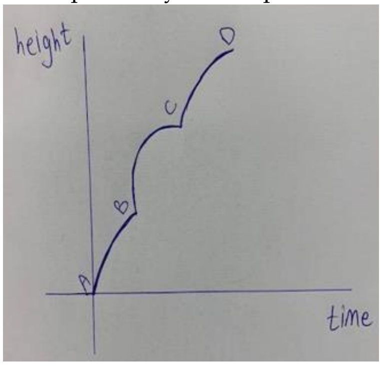
The goal of the study was to analyze the increase in sophistication of arguments over time (Schwarz & Baker, 2017). To this end, we focused on examining the level of appropriateness of the students' arguments in their written work in terms of three aspects: identifying variables, forming relations between them, and noticing contextual features (Ayalon et al., 2016a, 2016b). Each graph with its accompanying explanation (provided by individuals or by the group) was coded according to the rubric in **Table 2**.

Table 2. Rubric for coding the level of appropriateness of students' graphs

Rubric for coding
1. Inappropriate: evidence of graphing misconception
2. Partially appropriate; divided into three sub-categories: a. Choosing one correct and one incorrect variable and forming an irrelevant relation b. Choosing relevant variables but forming an irrelevant relation between them c. Failing to consider all contextual features
3. Appropriate

Table 3. Illustration of coding for category 1: Inappropriate: Evidence of graphing misconception

<table>
<tr><th>Category 1: Inappropriate: Evidence of graphing misconception</th><th>Illustrative evidence</th></tr>
<tr><td>The following graph was suggested by the group for the "filling a container" situation, accompanied by their explanation.



"The shape of the container, the wide part of the container–that means from point A to B–is slightly curved. The shape of the container, the narrow part of the container–that means from B to C–is more curved. The shape of the container, the wide part of the container–that means from C to D–is slightly curved."</td><td>The following conversation is taken from the transcript of the group interaction:
Anna: "What's that bulge?" (points to the shape of the graph)
Tom and Eva: "Because of the shape of the container. The container is curved and not straight so the graph must be curved at exactly the same way."
Eva: "The beginning and the end of the graph should be the same shape because of the shape of the container. The first part should be curved and the second part more curved and the third part as curved as the first part because the shape of the container is wide, narrow, wide."</td></tr>
</table>

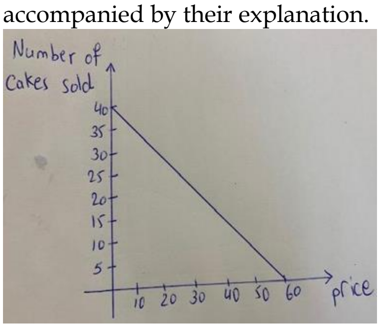
The analysis process also included transcripts of the group (videoed) interactions and individual interviews focused on students' discussions on variables, relations between variables, and contextual features to seek confirmation of the previous coding of the students' written products (Creswell, 2007). Decisions about interpretations were made collaboratively among the research team members through repeated cycles of check coding and discussion to reach consensus and refine the coding (Miles & Huberman, 1994).

Below we provide further information on the data-analysis process. For each of the five categories (1, 2a, 2b, 2c, and 3) we present an illustration of a work product that was assigned to the relevant category, together with supporting evidence taken from the transcripts of the group interactions and/or the individual interviews. **Table 3** presents an illustration for the category of *Inappropriate: graphing misconception*. In particular, arguments matching this category were characterized by a choice of one relevant variable with picture/graph confusion (Schultz et al., 1986). As illustrated in the example, the students interpreted the shape of the graph of the situation as a picture of that situation (the container). Note that we had to translate the axis labels of the graphs into English, so the graphs presented in the paper are reproductions of the students' drawings.

Table 4 presents an illustration for the category of *partially appropriate*. Three distinct sub-categories were

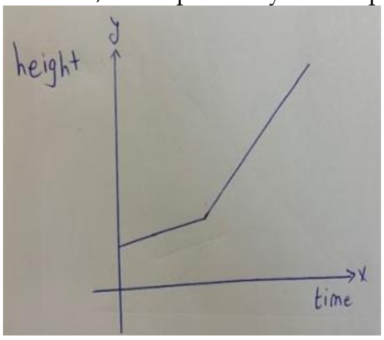
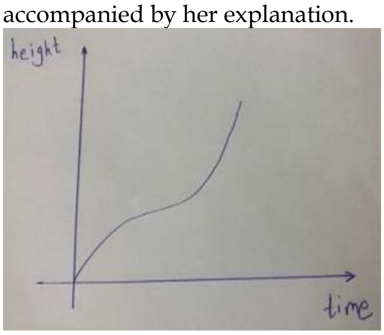
Table 4. Illustration of coding for category 2: Partially appropriate

<table>
<tr><th>Category 2: Partially appropriate</th><th>Illustrative evidence</th></tr>
<tr><td colspan="2">2a. One correct and one incorrect variable and irrelevant relation</td></tr>
<tr><td>The following graph was suggested by the group for the "pricing cakes" situation, accompanied by their explanation.



"We decided on the names of the axes according to the given situation … When the price is lower, more cakes will be bought, and when the price is higher, fewer cakes will be bought."</td><td>The following excerpt is taken from the transcript of the group interaction:
Tom and Liam: "As the price increases the number of cakes sold decreases."
The following excerpt is taken from an individual interview with Tom:
"The effect of price is on the number of cakes sold."</td></tr>
</table>

Table 4 (continued). Illustration of coding for category 2: Partially appropriate

Category 2: Partially appropriate	Illustrative evidence
2b. Relevant variables but irrelevant relation	
The following graph was suggested by the group for the "watering a plant" situation, accompanied by their explanation.  "We didn't begin from the zero point because there was a plant initially, and then we continued the function by increasing at a slow rate. Then, after Jack watered the plant, the function rose quickly."	The following excerpts are taken from the transcript of the group interaction: Tom: "I started with a constant rate." Anna: "The second part of the graph must be an increasing linear function."
2c. Failing to consider all contextual features	
The following graph was suggested by Roni for the "watering a plant" situation, accompanied by her explanation.  "Started from zero point, rose quickly and then slowly, and then rose more quickly."	The following excerpt is taken from the transcript of the group interaction: Roni: "The graph does not start from the zero point." The following excerpts are taken from individual interview with Roni: "The graph wasn't good because I started from the zero point and must start with a point on the axis."…"The rise at the beginning of the graph is wrong, because I started the graph before the situation where the seedling was watered."

identified. One is related to *choosing one correct and one incorrect variable and forming an irrelevant relation*; the second is related to *choosing relevant variables but forming an irrelevant relation between them*; and the third is related to *failing to consider all contextual features.*

As can be seen in the first example, the students chose one relevant variable (price) and one irrelevant variable (number of cakes sold) for describing the "pricing cakes" situation: "Karin, a cake shop owner, has to decide on the price at which to sell a new chocolate cake to customers. If the price is too low, she will lose money, but if the price is too high, fewer chocolate cakes will be sold and she will lose money. She needs to choose a reasonable price so that she can sell enough cakes to make a profit." The issue of profit, which is emphasized in the situation, was not chosen as a variable by the students. In the second example, the students chose relevant variables for the "watering a plant" situation (time and height): "Jack forgot to water his peas. The seedling dried up and its growth decelerated for a while. Then Jack remembered to water the sapling, and its growth accelerated." However incorrectly assumed linearity. In the third example, one student failed to consider all contextual features in the "watering a plant" situation.

Table 5 presents an illustration for the category of *appropriate*. As illustrated in the example, the student appropriately interpreted the "watering a plant" situation, choosing relevant variables, building correct relations between the variables, and attending to the contextual feature of not starting from the zero point of growth.

Analysis of students' evaluation of fictitious example-graphs

As described above, each cycle involved students in critiquing graphs together in groups, including writing the strength/s and the weakness/es of three fictitious example graphs (**Appendix A**, activity 3). The three graph examples were each designed to incorporate a combination of different weaknesses and strengths in terms of the three aspects: the choice of variables, the relation between them, and contextual features of the real-life situation. We asked students to critique the three fictitious graphs after constructing their own graph. Then, after evaluating the fictitious students' graphs, we invited the students to revisit their own constructed group graph and decide whether to revise it.

The analysis of the students' evaluation of the three fictitious example graphs in the group focused on the students' attention to the strengths and weaknesses of

Table 5. Illustration of coding for category 3: Appropriate

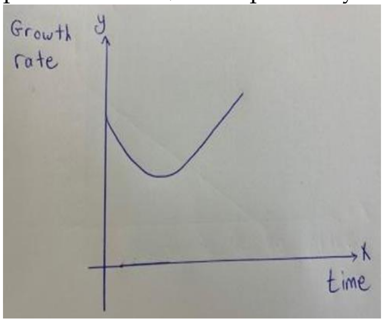
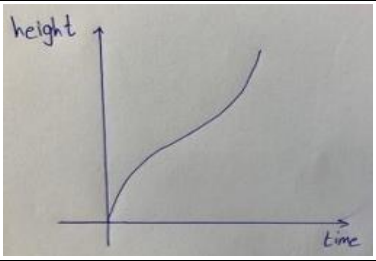
Category 3: Appropriate	Illustrative evidence
The following graph was suggested by Anna for the "watering a plant" situation, accompanied by her explanation.  "The graph started from a certain point, decreased and then increased gradually."	The following excerpts are taken from the transcript of the group interaction: Anna: "I started from a certain point and because he stopped watering, its growth was slower; then went down, and then he watered it again, so its growth accelerated." Anna: "It is like, I think that this graph [points to another graph suggested by Eva, which presented a linear relation between time and growth rate] is too sharp. It did not decrease in one shot. It is like... it was gradual." The following excerpt is taken from an individual interview with Anna: "I think that the beginning is not from the zero point because it was written that he [already] had a seedling and forgot to water it."

Table 6. The group's evaluation of one of the fictitious example-graphs given in the situation "filling a container"

The given graph	Strengths	Weaknesses
	The labels of the axes (height/time)	The beginning and the end are not the same shape

each graph. The aim of the analysis was to provide us with more information to help us interpret any modification made by the group in their final graph compared to their initial one. For example, **Table 6** presents the group's evaluation of one of the fictitious example-graphs given in the situation "filling a container". The strength of the given graph is in the choice of variables (height/time), and the weakness of the graph is related to the change of the height over time (which should first increase at a fast rate and then at a slower rate).

As seen in **Table 6**, the students correctly identified the strength of the graph. However, they could not see the weakness in the graph in terms of the relation between the variables. Instead, they wrote that "the beginning and the end are not the same shape." This seemed related to the discussion they had held before the critiquing activity, in which they emphasized that the beginning and the end of the graph should be the same shape. Since the students had divided the container into three parts, they drew a graph corresponding to those three parts, without paying attention to the differences between the parts of the graph in terms of their convexity/concavity.

FINDINGS

To share our analysis of the appropriateness of the students' arguments throughout participation in a sequence of collaborative activities, we present our findings chronologically according to the three activity cycles. Three tables each present the categorization of the group's six individual graphs constructed in the first phase of every cycle[1], the group's initial graph before critiquing the fictitious examples-graphs and the group's final graph after critiquing the examples.

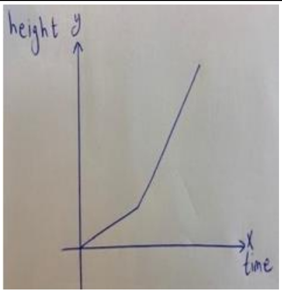
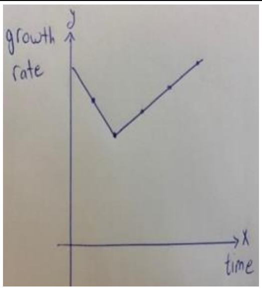
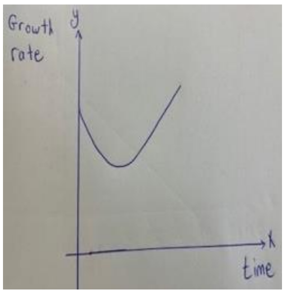
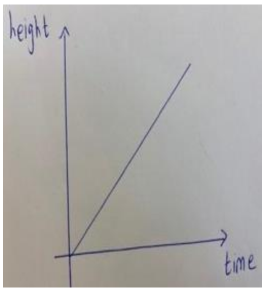
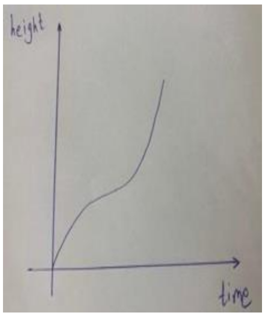
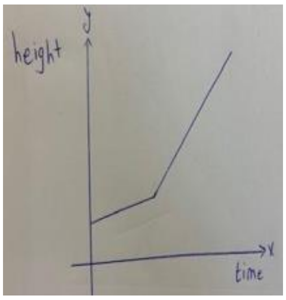
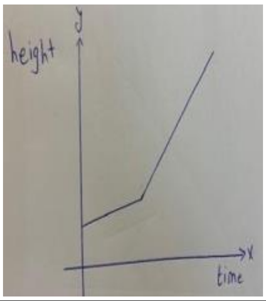
Students' Responses During the 'Watering a Plant' Sequence

Table 7 presents the findings related to situation 1 "watering a plant": "Jack forgot to water his peas. The seedling dried up and its growth decelerated for a while. Then Jack remembered to water the sapling, and its growth accelerated."

As shown in **Table 7**, all individual graphs related to situation 1, except for two, were coded as category 2b: Choosing relevant variables but forming an irrelevant relation between them. One graph, provided by Roni, was coded as category 2c: Failing to consider all contextual features; and one graph, provided by Anna, was coded as category 3: Appropriate. All products categorized as category 2b were characterized by assuming a linear relation between variables. For example, Tom chose the variables height/time but incorrectly drew a linear relation between them. Omer also chose the variables of height/time; however, not only did he mistakenly assume linearity, but he did not

[1]In the third cycle, only five students participated, as Roni was absent from school that day.

Table 7. Constructed individual and group graphs for situation 1, "watering a plant"

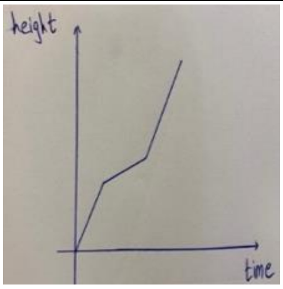
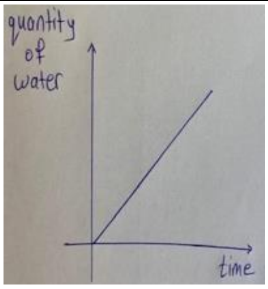
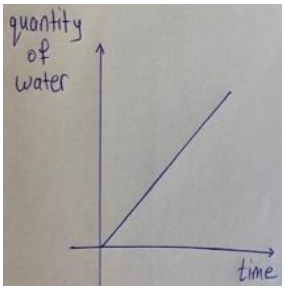
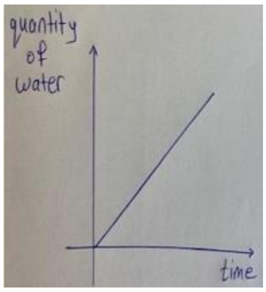
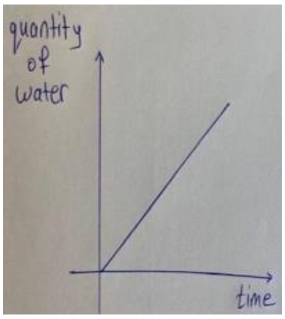
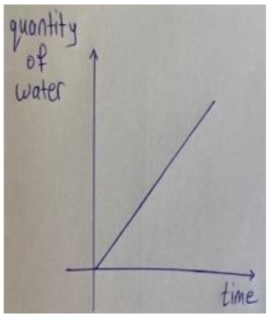
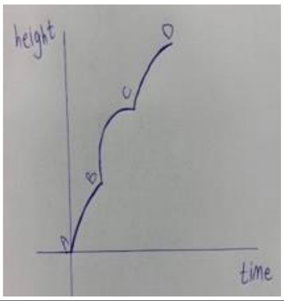
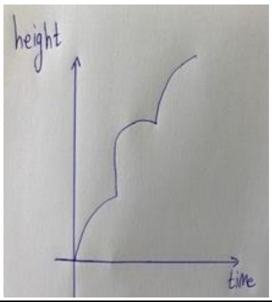
Graph constructor	The proposed graph	Code	Graph constructor	The proposed graph	Code
Tom		2b	Eva		2b
Anna		3	Omer		2b
Liam		2b	Roni		2c
Initial group graph		2b	Final group graph		2b

Note. Although only the graph is presented in the table, the categorization is based on analysis of multiple data sources–written justifications, interviews, and group interactions, as was illustrated in the data analysis section

attend to the change of rate within the situation. In contrast to these students, Roni constructed a correct relation between the variables of height and time. Still, she started the graph at the origin, not considering that the height of the plant was not zero at the beginning of the situation. In the interview, he acknowledged his mistake, saying: "The graph wasn't good because I started from the zero point … I started the graph at a time that precedes the story in the situation."

Anna, the only one who constructed a valid graph, attempted to explain it to her peers in the group discussion: "It started from a certain point and then, because Jack stopped watering the plant, its growth was slower, and then he watered it again, so its growth accelerated … It is like, I think that this graph [points to another graph suggested by Eva, which presented a linear relation between time and growth rate] is too sharp. It did not decrease in one shot. It is like, it was gradual." In the interview, she added: "I think that the beginning is not at the zero point because we were told that Jack at some point forgot to water the plant." From the transcripts of group interactions, we were able to learn that Anna, who provided an appropriate graph, tried to convince her peers about the correctness of her graph, but did not succeed. They rejected her idea, probably because they continued to think about the relation between height and time and not between the growth rate and time. For example, one of them said: "How can it be that the height went down and then up!" Tom, who constructed a linear relation between height and time, succeeded in convincing the other group members that the graph must rise slowly and then quickly. At the end of the discussion, Anna was also

Table 8. Constructed individual and group graphs for situation 2, "filling a container"

Graph constructor	The proposed graph	Code	Graph constructor	The proposed graph	Code
Tom		2b	Eva		2a
Anna		2a	Omer		2a
Liam		2a	Roni		2a
Initial group graph		1	Final group graph		1

Note. Although only the graph is presented in the table, the categorization is based on analysis of multiple data sources–written justifications, interviews, and group interactions, as was illustrated in the data analysis section

convinced, and she said: "It makes sense to me." The students decided to follow Tom's idea, but not from the zero point.

Interestingly, in part 3 of the group evaluation of the given example-graphs, the students Anna and Roni positively evaluated one of the functions as a nonlinear function and negatively evaluated another function, which was a linear function. All the students emphasized that the height cannot decrease, and that the beginning of the graph is not from the zero point. Yet in part 4, when asked to revisit their group graph and consider modification, the students chose to keep their linear graph. They said that they were confident about their graph, saying: "There is no need to change the graph" and "It makes sense to me" and "We are confident about our graph."

Figure 2. Container-2 (Ayalon et al., 2021)

Students' Responses During the 'Filling a Container' Sequence

Table 8 presents the findings related to situation 2 "filling a container": A tap is turned on hard, and water rushes into the container in **Figure 2** at a constant rate.

As shown in **Table 8**, most of the individual written products related to situation 2 were coded as category

2a: choosing one correct and one incorrect variable and forming an irrelevant relation. Five students treated the quantity of water as being dependent on time, instead of the height of the water being dependent on time. Two possible explanations may be drawn from the transcripts of the group interactions and interviews regarding the students' tendency to choose "quantity of water" instead of its "height". One possible explanation may be that some of the students did not notice the picture given in the situation, as Liam said in his interview: "I didn't notice the picture; I just looked at what was written and saw that it said it was pulling water at a constant rate." Another optional explanation for choosing "quantity of water" instead of "height" is that they noticed the picture but did not take the shape of the container into consideration. For example, during the group discussion Anna said: "You do not have to refer to the shape of the container; it is the quantity of water that is at a constant rate." As also shown in **Table 8**, Tom was the only one whose product was coded as category 2b: Choosing relevant variables but forming an irrelevant relation between them. This student referred to the shape of the container and chose height and time as variables but failed to formulate the correct relation between them by assuming a linear relation and did not attend to the change in the rate within the situation.

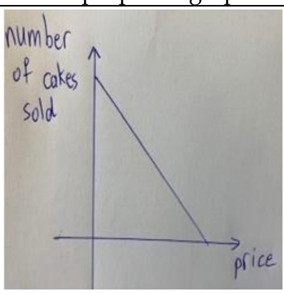
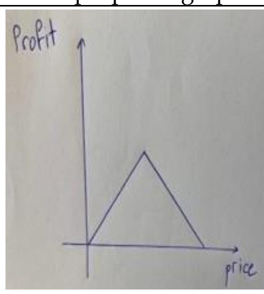
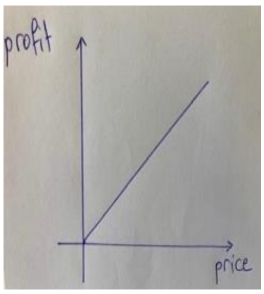
Interestingly, the graph constructed by the whole group was different from the ones proposed by each of them individually. Both group's graphs–initial and final–were characterized by picture/graph confusion (category 1) and followed Tom's [mistaken] ideas. During the group's discussion, Tom changed his mind concerning the graph he had suggested previously and said: "Now I have noticed something … the narrow part of the container is filled faster … and no section of the container is straight." He drew a graph that imitates the shape of the container, saying: "The shape of the container is curved, so the graph must be curved at exactly the same way." The group decided that "The first part is slightly curved, the second part is more curved, and the third part is slightly curved, similar to the shape of the container … and the first part and the third part are the same [curved] shape in the container, so it has to be the same shape in the graph." Then, the students evaluated the three example graphs, emphasizing that the height cannot decrease, the beginning and the end of the graph must be the same shape, and the graph should describe the shape of the container which starts out wide, then becomes narrow and in the end is wide [again]. They decided not to change their initial group graph.

Students' Responses During the 'Pricing Cakes' Sequence

Table 9 presents the findings related to situation 3 "pricing cakes": Karin, a cake shop owner, has to decide on the price at which to sell a new chocolate cake to customers. If the price is too low, she will lose money, but if the price is too high, fewer chocolate cakes will be sold and she will lose money. She needs to choose a reasonable price so that she can sell enough cakes to make a profit.

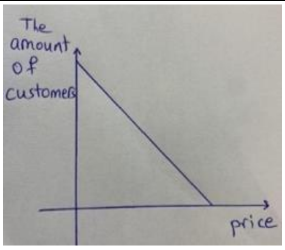
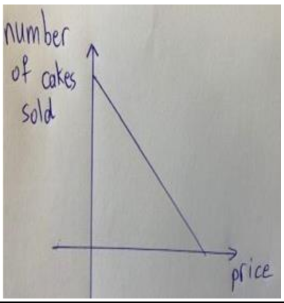
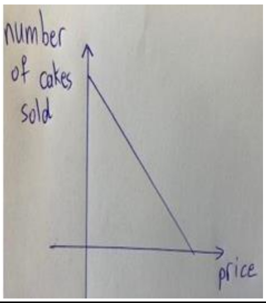
As shown in **Table 9**, two of the individual written products related to situation 3 were coded as category 2a (choosing one correct and one incorrect variable and forming an irrelevant relation). These students used as the dependent variable "number of cakes sold" instead of "profit" and considered a decreasing linear relation between price and number of sold cakes. While Liam explained that the higher the price of a cake, the lower the number of customers who will buy the cakes, Tom

Table 9. Constructed individual and group graphs for situation 3, "pricing cakes"

Graph constructor	The proposed graph	Code	Graph constructor	The proposed graph	Code
Tom		2a	Eva		2b
Anna		2b	Omer		2b

Note. Although only the graph is presented in the table, the categorization is based on analysis of multiple data sources–written justifications, interviews, and group interactions, as was illustrated in the data analysis section

Table 9 (continued). Constructed individual and group graphs for situation 3, "pricing cakes"

Graph constructor	The proposed graph	Code	Graph constructor	The proposed graph	Code
Liam		2a	Roni	Absent from class	NA
Initial group graph		2a	Final group graph		2a

Note. Although only the graph is presented in the table, the categorization is based on analysis of multiple data sources–written justifications, interviews, and group interactions, as was illustrated in the data analysis section

explained that as the price increases the number of cakes sold decreases. Three other individual written products were coded as category 2b (choosing relevant variables but forming an irrelevant relation between them). These students correctly chose price and profit but formed an irrelevant relation between these variables, assuming a linear relation. Anna and Omer drew linear functions, but neither of them attended to the change in the rate within the situation. Anna drew a decreasing relation and explained in the group interaction: "When the price is high, then a small number of cakes is bought so the profit is little; and when the price is low, then a larger number of cakes is bought, and the profit is big; and when the price is reasonable then a large number of cakes is bought, and the profit is reasonable." Omer drew an increasing linear function, identifying the height of the price with the height of the profit: "The price of cakes is the profit that she earns." Similar to Anna and Omer, Eva also drew a linear relation between profit and time. However, in contrast to those students, Eva drew a graph that starts out with an increasing linear function and then turns into a decreasing linear function. This student, although assuming linearity, attended to the complex nature of the profit variable. During the group interaction, she attempted to justify her graph: "When the price of the cake is low then the profit is not high, and when the price of the cake is higher, then the profit increases, but when the price is really high, the profit decreases again."

As also shown in **Table 9**, in both group's graphs, initial and final, the students chose the irrelevant variable of "number of cakes sold", although some of them individually chose the relevant variable of "profit". The transcript of the group interactions helped us make sense of reasons for the group's choice. The group members decided to follow Tom's idea. Tom presented his graph and explained: "Profit and the number of cakes sold are the same data; the profit is from the cakes that were sold… the graph describes the profit, and there is no need to have an axis with profit." There was a long discussion among the students, with several differing opinions, but they had a hard time reaching consensus. Eventually, they decided to draw Tom's graph as their group graph, although they were not convinced of its correctness. For example, Eva said: "I understood it, but it's not correct … I think that no one is right, but we will draw Tom's graph."

During the group evaluation of the three example graphs, the students correctly identified the variables profit/price as correct. They then tried to construct a graph for these variables but did not succeed. Eventually, they decided to keep the graph they had constructed before the evaluation (Tom's graph), despite recognizing that it was only partially correct.

In the interviews, some of the students claimed that this situation was the most difficult. They said that the group members spent a lot of time discussing the different options of the three variables: price, profit and number of cakes, and that it was hard for them to construct an appropriate graph.

Overall, across the three situations, most of the students' arguments were coded as category 2a (choosing one correct and one incorrect variable and forming an irrelevant relation) or 2b (choosing relevant variables but forming an irrelevant relation between them). Two arguments were coded as category 1 (inappropriate: graphing misconception), one as category 2c (failing to consider all contextual features) and one as category 3 (appropriate). Many of the graphs also incorrectly assumed linearity. Additionally, the students did not change their initial group graph, which was drawn before critiquing the fictitious examples. As an exception, in the third cycle, following the critique, the students realized that the correct variables should be price and profit (and not number of cakes sold, as was their original choice). They tried to construct a graph for the variables of price and profit but did not succeed and, therefore, decided to keep their graph as it was before the critique. Finally, we found that across the situations, after having rich and lengthy discussion, the group nevertheless repeatedly defaulted to using or adapt Tom's ideas for their group graphs, even though in one situation another group member had drawn the correct (individual) graph. In situation 2 and situation 3, the group chose to copy Tom's graphs before and after the examples critique. In situation 1, the group adapted Tom's graph. During his final interview, Tom reflected, "I feel like Albert Einstein".

DISCUSSION AND CONCLUSION

In this study, we examined the appropriateness of a group of 9th grade students' arguments through their participation in a sequence of collaborative argumentation activities focused on constructing and critiquing graphs representing real-life functional situations. We used the Toulminian argument schema in the student handouts, incorporated students' critiques of fictitious example graphs, and developed collaborative argumentation learning tasks to support students in building criteria for the construction and evaluation of arguments in order to encourage their productivity in mathematical argumentation.

The analysis of student-produced graphs revealed that most individual graphs fell into two categories: selecting correct variables but forming irrelevant relationships (category 2b) or choosing one correct and one incorrect variable (category 2a). While students were initially successful in identifying variables such as time and height, many struggled to construct accurate relationships between them, often defaulting to linear assumptions, a misconception commonly documented in previous research (Janvier, 1981; Leinhardt et al. 1990; Thompson, 1994). Our findings suggest that this issue may have been influenced by the task design, its classroom implementation, and social dynamics within the group, although further research is needed to confirm these observations.

Across the three cycles, we observed that students tended to base their final group graphs on the ideas of one dominant figure, typically Tom, whose ideas were not always correct. In situation 1, the group graph was categorized as 2b, i.e., choosing relevant variables but forming an irrelevant relation between them. The group incorrectly assumed linearity but did account for contextual constraints in not starting their graph from the origin. It is worth noting that one student (Anna) had in fact constructed an appropriate graph for Situation 1 and another student (Roni) also constructed a nearly correct graph. This demonstrates that this functional situation was within the reach of some students at this grade level, but that other influences on students' decision making were involved. In situation 2, the group graph was coded as 1, and evidenced picture/graph confusion (linked to the shape of the container), well-known in the literature on student difficulties with graphs (e.g., Leinhardt et al., 1990; Schultz et al., 1986). This graph was different from all of the individual graphs; Tom had directed the process for its construction. In situation 3, the group graph was categorized as 2a, in containing an inappropriate variable. Despite acknowledging its imperfections, the group chose Tom's graph, suggesting a reliance on his authority rather than critical evaluation of the task.

While our study did not focus specifically on the social interactions that contributed to or hindered argumentation, these dynamics warrant further exploration. The teacher in her interview revealed that Tom was perceived as an authority by his peers: "The students have known Tom for many years, and they know that he usually gives correct answers." This finding aligns with existing research on peer authority in group work (Bucholtz & Hall, 2005). However, our study indicates that this reliance on a single student's authority can sometimes impede the group's ability to reach accurate mathematical conclusions, a point that deserves closer attention in future research.

Contrary to our expectations, we did not observe significant improvement in students' arguments across the iterative cycles of tasks. We speculate that the increasing complexity of the tasks may have amplified the students' difficulties in interpreting functional concepts through graphical representations. Additionally, the specific challenges varied by situation. In situation 1, the primary difficulty was assuming a linear relationship between variables, a common issue documented in research (Janvier, 1981; Leinhardt et al., 1990; Thompson, 1994). In situation 2, students struggled with variable selection, often focusing on irrelevant variables like the 'quantity of water' instead of the rate of pooling water, despite the task's explicit focus on rate. This confusion may stem from everyday experiences where students think about quantity rather than rate (Goldenberg & Kliman, 1988). In situation 3, difficulties were related to both identifying the variables–with a tendency to use an irrelevant variable 'number of cakes'–and forming a relation that did not match the contextual details, again, assuming linearity. A possible explanation for the difficulties encountered may be related to the multiple variables appearing in the narrative of this particular situation (price, number of cakes sold, and profit) where the choice is not trivial (Ayalon et al., 2016a). In the interviews, the students talked about the group members spending a lot of time discussing the various options of using two of the three variables: price, profit, and number of cakes. They said that it was difficult for them to construct an appropriate graph. Some students speculated that the 'number of cakes sold' and 'profit' are actually the same variable, since the profit derives from the number of cakes sold. Such an identification disregards the situation where not enough cakes are sold to make a profit. It is possible that the 'number of cakes sold' is more tangible to students from their own everyday life than 'profit'. Some students did use 'profit' as their dependent variable but nevertheless constructed an incorrect relation between the variables price and profit. This may result from the complex and unfamiliar nature of 'profit', which made it difficult for them to identify the correct direction of the graph (Ayalon et al., 2016a).

Our study was limited in that it primarily focused on the appropriateness of students' arguments rather than the full range of interactions and processes involved in their argumentation. Although we did not systematically investigate the impact of specific task characteristics (e.g., comparing graphs, writing down differences, and reaching consensus), our analysis of student discourse hints at potential factors that both enabled and inhibited argumentation. For example, we included a peer critique activity in each cycle, where students critiqued fictitious example graphs, aiming to foster their ability to evaluate arguments and revise their own graphs. However, in all cycles, the students chose not to revise their group graphs despite their critiques. This contrasts with literature that advocates for peer critique as a learning tool (Erduran & Jiménez-Aleixandre, 2008). Our findings suggest that peer critique may be less effective when students lack a deep understanding of the mathematical concepts at hand, particularly in real-life situations that may be unfamiliar. It is also possible that students' lack of prior experience with critiquing graphs, as noted in interviews with both the teacher and the students, limited their ability to use peer critique effectively. There was some hint, in situation 3, that evaluation of the three fictitious example-graphs did lead students to reconsider the variables they had chosen in their initial group graph: 'price of a cake' and 'number of cakes sold'. They realized that 'profit' better fits the situation than 'number of cakes sold', and they made a serious effort together to construct the graph correctly, but with no success. This suggests that the evaluation of example graphs did have some positive impact, but that the impact was not fully realized.

Parts of Toulmin's (1958) schema (see **Figure 1**) was used as a means for structuring the task in which the students were engaged and to provide a means for structuring students' discussions (e.g., Chin & Osborne, 2010). In their written products, the students were found to distinguish appropriately between data (*a specific piece of evidence from the story of the situation*) and warrants (*a mathematical reason why my graph matches the piece of evidence*) when justifying their drawn graphs. In addition, the interactions among students were characterized by their presenting their ideas in terms of data extracted from the story situations and warrants in terms of the mathematical reasons for the graph being compatible with the given situation. The students provided justifications and also asked peers to provide justifications for their point of view. It is possible then, as was found in several studies in science education, that Toulmin's (1958) schema helped to structure the students' group discussions in our study. While this structured approach may have supported some aspects of their argumentation, further research is needed to explore how Toulmin's (1958) framework, commonly used in science education (Chin & Osborne, 2010), can be adapted more effectively for mathematics education (Cavagnetto & Kurtz, 2016).

In conclusion, this study offers preliminary insights into the factors that may influence the appropriateness of students' arguments in collaborative mathematical tasks. The findings point to the importance of managing both social dynamics and task complexity to support productive argumentation. While Toulmin's (1958) schema shows promise in structuring mathematical discussions, our results suggest that additional scaffolding may be necessary to ensure student success. Finally, the role of peer critique in improving argumentation remains uncertain when students are working with unfamiliar mathematical concepts. These findings contribute modestly to the broader discussions on how to structure collaborative mathematical tasks and support student learning in group settings.

Author contributions: MA, KJW, & RS: methodology, analysis, writing - original draft; **MA & KJW:** conceptualization & resources, writing - review & editing; **MA & RS:** investigation and project administration. All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was approved by the University of Haifa Human Research Ethics Committee on 4 November 2018 with registration number 312/18. Only students whose parents have given their written consent participated in the study. The authors further stated that the researchers provided the management of the educational institution with the letters for parents on the subject. The management arranged the distribution of the letter to parents through the students. The parents who agreed to their child's participation in the study returned the written consent signed by them to the management of the educational institution, from which the researchers collected the aforementioned forms.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Andriessen, J. E. B., & Schwarz, B. B. (2009). Argumentative design. In N. Muller-Mirza, & A.-N. Perret-Clermont (Eds.), *Argumentation and* *education–Theoretical foundations and practices* (pp. 145-174). Springer. https://doi.org/10.1007/978-0-387-98125-3_6

Asterhan, C. S. C., & Schwarz, B. B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist, 51*(2), 164-187. https://doi.org/10.1080/00461520.2016.1155458

Ayalon, M., & Even, R. (2016). Factors shaping students' opportunities to engage in argumentative activity. *International Journal of Science and Mathematics education*, *14*(3), 575-601. https://doi.org/10.1007/s10763-014-9584-3

Ayalon, M., Watson, A., & Lerman, S. (2016a). Progression towards functions: Students' performance on three tasks about variables from grades 7 to 12. *International Journal of Science and Mathematics Education* *14*(6), 1153-1173. https://doi.org/10.1007/s10763-014-9611-4

Ayalon, M., Watson, A., & Lerman, S. (2016b). Reasoning about variables in 11 to 18 year olds: Informal, schooled and formal expression in learning about functions. *Mathematics Education Research Journal, 28*, 379-404. https://doi.org/10.1007/s13394-016-0171-5

Ayalon, M., Watson, A., & Lerman, S. (2018). Comparison of students' understanding of functions in classes following English and Israeli national curricula. *Educational Studies in Mathematics, 97(3),* 255-272. https://doi.org/10.1007/s10649-017-9798-8

Ayalon, M., Wilkie, K. J., & Eid, K. H. (2021). Relating students' emotions during argumentative discourse to their learning of real-life functional situations. *Educational Studies in Mathematics, 110*, 23-48. https://doi.org/10.1007/s10649-021-10121-5

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies, 7*(4-5), 585-614. https://doi.org/10.1177/1461445605054407

Burkhardt, H., & Swan, M. (2012). Designing assessment of performance in mathematics. *Educational Designer*, *2*(5), 1-40.

Cavagnetto, A. R., & Kurtz, K. J. (2016). Promoting students' attention to argumentative reasoning patterns. *Science Education, 100*(4), 625-644. https://doi.org/10.1002/sce.21220

CCSSI. 2010. Common core state standards for mathematics. *National Governors Association Center for Best Practices and the Council of Chief State School Officers*. http://www.corestandards.org/the-standards

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*(4), 359-387. https://doi.org/10.1007/BF01273371

Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of Research in Science Teaching, 47*(7), 883-908. https://doi.org/10.1002/tea.20385

Chua, B. I., (2016). Justification in Singapore secondary mathematics. In P. C. Toh, & B. Kaur (Eds.), *Developing 21st century competencies in the mathematics classroom* (pp. 165-188). World Scientific. https://doi.org/10.1142/9789813143623_0010

Clement, J. (1985). Misconceptions in graphing. In *Proceedings of the 9th International Conference for the Psychology of Mathematics Education* (Vol. 1, pp. 369-375). Utrecht University.

Corrbet, D., & Wilson, B. (1995). Make a difference with, not for, student. *Educational Researcher, 24*, 12-17. https://doi.org/10.3102/0013189X024005012

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). SAGE.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction, 20*, 399-483. https://doi.org/10.1207/S1532690XCI2004_1

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in science education. Perspectives from classroom-based research*. Springer. https://doi.org/10.1007/978-1-4020-6670-2

Felton, M., Garcia-Mila, M., & Gilabert, S. (2009). Deliberation versus dispute: The impact of argumentative discourse goals on learning and reasoning in the science classroom. *Informal Logic, 29*(4), 417-446. https://doi.org/10.22329/il.v29i4.2907

Francisco, J. M., & Maher, C. A. (2005). Conditions for promoting reasoning in problem solving: Insights from a longitudinal study. *The Journal of Mathematical Behavior,* 24(3-4), 361-372. https://doi.org/10.1016/j.jmathb.2005.09.001

Goldenberg, E. P. (1987). Believing is seeing: How preconceptions influence the perceptions of graphs. In J. Bergeron, N. Herscovits, & C. Kieran (Eds.), *Proceedings of the eleventh conference of the International Group for the Psychology of Mathematics Education* (pp. 197-203).

Goldenberg, E. P., & Kliman, M. (1988). *Metaphors for understanding graphs: What you see is what you see.* https://eric.ed.gov/?id=ED303369

Janvier, C. (1981). Use of situations in mathematics education. *Educational Studies in Mathematics, 12*(1), 113-122. https://doi.org/10.1007/BF00386049

Jimenez-Aleixandre, M. P., &Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran, & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3-27). Springer. https://doi.org/10.1007/978-1-4020-6670-2_1

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*(1), 1-64. https://doi.org/10.3102/00346543060001001

Lemke, J. (1990). *Talking science: Language, learning and values*. Ablex.

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett, & P. Shah (Eds.), *Proceedings of the 33rd Carnegie Symposium on Cognition*. Lawrence Erlbaum Associates, Inc.

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229. https://doi.org/10.1002/sce.20364

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. SAGE.

Ministry of Education. (2020). *The adult character skills document*. https://meyda.education.gov.il/files/Mazkirut_Pedagogit/MadaTechnologya/yesodi/boger2030.pdf

Ministry of Education. (2024). *The new mathematics curriculum for middle-high school*. https://pop.education.gov.il/tchumey_daat/matmatika/chativat-beynayim/teaching-mathematics/tohnit-limudim/

Mueller, M., Yankelewitz, D., & Maher, C. (2014). Teachers promoting student mathematical reasoning. *Investigations in Mathematics Learning, 7*(2), 1-20. https://doi.org/10.1080/24727466.2014.11790339

Reuter, F. (2023). Explorative mathematical argumentation: A theoretical framework for identifying and analysing argumentation processes in early mathematics learning. *Educational Studies in Mathematics, 112*, 415-435. https://doi.org/10.1007/s10649-022-10199-5

Schultz, K., Clement, J., & Mokros, J. (1986). *Adolescents' graphing skills: A descriptive analysis* [Paper presentation]. The Meeting of the American Educational Research Association.

Schwarz, B. B., & Baker, M. J. (2017). *Dialogue: Argumentation and education: History, theory and practice*. Cambridge University Press. https://doi.org/10.1017/9781316493960

Slavit, D. (1997). An alternate route to the reification of function. *Educational Studies in Mathematics, 33*(3), 259-281. https://doi.org/10.1023/A:1002937032215

Staples, M. (2014). Supporting student justification in middle school mathematics classrooms: Teachers' work to create a context for justification. *CRME Publications*. 4. https://digitalcommons.lib.uconn.edu/merg_docs/4

Thompson, P. W. (1994). Images of rate and operational understanding of the fundamental theorem of calculus. *Educational Studies in Mathematics, 26*(2-3), 229-274. https://doi.org/10.1007/BF01273664

Toh, P. C., & Kaur, B. (Eds.). (2016). *Developing 21st century competencies in the mathematics classroom*. World Scientific. https://doi.org/10.1142/10132

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.

van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616389

Weber, K., Maher, C., Powell, A., & Lee, H. S. (2008). Learning opportunities from group discussions: Warrants become the objects of debate. *Educational Studies in Mathematics, 68*(3), 247-261. https://doi.org/10.1007/s10649-008-9114-8

Wood, T., Williams, G., & McNeal, B. (2006). Children's mathematical thinking in different classroom cultures. *Journal for Research in Mathematics Education, 37*(3), 222-255.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458-477. https://doi.org/10.5951/jresematheduc.27.4.0458

Zambak, V. S., & Magiera, M. T. (2020). Supporting grades 1-8 pre-service teachers' argumentation skills: Constructing mathematical arguments in situations that facilitate analyzing cases. *International Journal of Mathematical Education in Science and Technology, 51*(8), 1196-1223. https://doi.org/10.1080/0020739X.2020.1762938

APPENDIX A: SAMPLE ARGUMENTATION TASKS: CYCLE #3

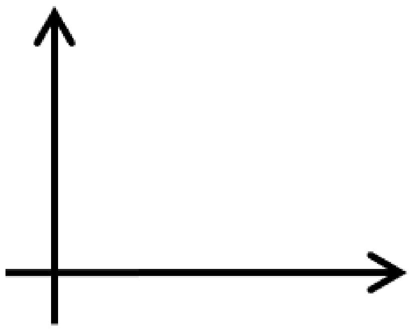
(Handouts have been condensed.)

Activity 1: Individual Graph Construction

A cake shop owner has to decide on the price at which to sell a new chocolate cake to customers. If the price is too low, she will lose money, but if it is too high, fewer chocolate cakes will be sold, and the owner will also lose money. She has to choose a moderate price so that she can sell enough cakes to make a profit.

Draw a graph that you think best matches the situation.

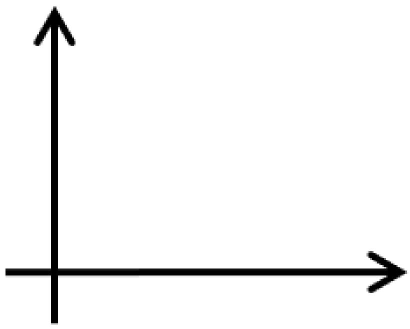
Label each axis and other key features of your graph to explain how they match the description of the situation.



Activity 2: Compare Individual Graphs in a Group; Write Down Differences; Draw a Group Graph

We have compared our individual graphs of the situation.

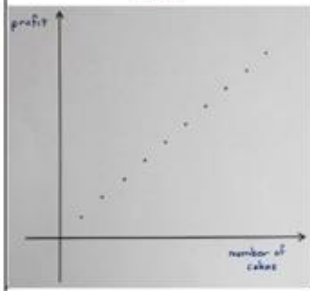
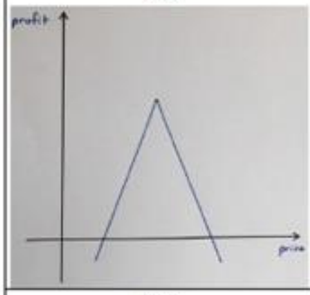
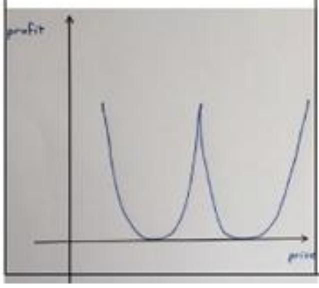
1. List the differences between the graphs drawn by your group (explain):
2. For each difference, describe how you reached consensus:
3. **Claim**: We have decided that the following graph matches the situation.



Activity 3: Group Critique of Fictitious Student Ideas

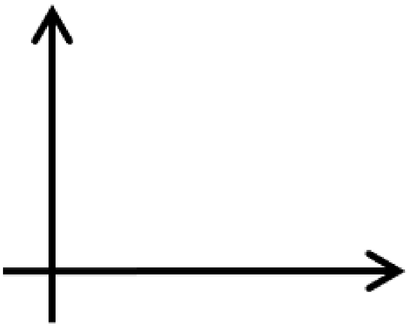
Here are three students' drawings of their graph to match the same situation. Critically assess each student's graph in terms of how realistically it matches the situation.

Data (specific piece of evidence from the story of the situation)	Warrant (mathematical reason why the graph matches the piece of evidence)

ALEX	Strength/s	Weakness/es
		
BEN	Strength/s	Weakness/es
		
CATH	Strength/s	Weakness/es
		

Activity 4: Reach Agreement on a Final Group Graph With a List of Data and Warrants

1. Re-visit your group's graph and decide if you would like to change it (draw below). Justify why you think your original or your revised graph matches the situation.



2. Justification for our final graph (initial or revised):

   We are keeping/changing our graph because …

3. What is a possible feature in your graph that other students might question or disagree with?
   How might you respond to them?

APPENDIX B: POST-CYCLE INTERVIEW: FOCUS GROUP AND TEACHER

Student

1. How did you decide on your own graph initially? (Look at the handout together.)
2. What differences were there between your own initial graph and the initial graph of your group? Were you happy with the group decision for this graph?
3. What differences were there between your initial and final group graphs? Were the three student example-graphs helpful for making your final group graph? Why/why not? Were you happy with the group decision for the final graph?
4. Did anything surprise you about the different group graphs presented to the whole class?
5. I noticed that during the lesson, you ………… (event observed). Can you explain to me what was happening from your perspective?
6. I see you wrote that you learnt …………... (last two questions on written reflection handout).
7. Is there anything else you would like to mention?

Teacher

1. What do you remember as a 'stand-out' from the two lessons about (the story situation)?
2. Did you experience any surprising or interesting or difficult moments during the lessons?
3. What do you think about the engagement of the class during the cycle (was it different or similar to usual for this class)?
4. What did you notice about the students' responses to the tasks?
5. In the first/second lesson in the cycle, I noticed that………….. (critical event noticed by researcher). Can you explain what you noticed from your perspective?

After Third Cycle

1. What do you think about the students' progress in learning with these tasks over the whole sequence?
2. If you repeated the sequence, what do you think you would do differently (give reason)? Grouping? Discussions? Sequencing? Tasks?