Mathematics Teachers' Use of Content-specific Data for Dialogic Grouping

Rotem Abdu
David Yellin Academic College of Education

Shai Olsher
University of Haifa

Received: February 2024 II Accepted: February 2025

Group composition affects learning by individuals. Dialogic pedagogy approaches demonstrate that this is particularly true when each grouped student knows something others do not (i.e., *mutuality* grouping). Learning analytics can help grouping by providing teachers with data on students' content-specific learning. What are mathematics teachers' considerations in grouping students based on such data? We analysed fifty-three acts of grouping by nine mathematics teachers, who used data about students' solutions to a mathematical task on linear functions to group students into pairs. We propose a schematic model including two types of considerations: interpersonal (here, *mutuality, encompassing*, and *similarity*) and content-specific (here, *methods of construction*, *function orientation,* and *correctness*). In this study, encompassing was the leading interpersonal characteristic, and function orientation was the leading content-specific characteristic. Moreover, different teachers formed the same groups using various considerations. The teachers utilised learning analytics to group students, and we modelled their grouping considerations.

Keywords: mathematics teacher education research • group learning • personal example space • teachers' grouping considerations

Introduction

This study queries teachers' considerations to forming dyads for small group learning based on students' examples of the concept of linear functions. Group learning (also known as collaborative or cooperative learning) happens when two or more people try to learn something together (Dillenbourg, 1999). Group learning is usually considered an effective methodology at the individual and interpersonal levels (Slavin, 1980). For example, group members may have a shared goal, such as solving a mathematical problem. In contrast, individuals may have different learning goals in other group learning situations, such as learning from each other. There are, however, many forms of group learning but not all are effective (Barron, 2003; Schwarz et al., 2000). To make group learning effective, students must use certain behaviours, such as engaging with each other, encouraging group members to participate, and actively listening to the ideas of group members (Abdu et al., 2015; Cohen, 1994; Schwarz et al., 2015; Webb, 2009; Webb et al., 2014). The effectiveness of group learning is also affected by the individuals of the group: who they are, what they know, and with whom they are learning. Allocating students to groups for learning (grouping) requires consideration of contextual, personal, and interpersonal learning criteria (Azmitia & Montgomery, 1993; Maqtary et al., 2019; Webb, 2009).

The context of the present study is contemporary technological developments of automated formative assessment, in which data on students' (mathematical) thinking and learning is collected and analysed to provide instant fine-grained personalised feedback (Hopfenbeck et al., 2023; Olsher et al., 2016). The goal of the present study is to study mathematics teachers' considerations of grouping students into pairs using such personalised feedback to enhance individual learning.

Grouping

Many factors govern how groups are formed. Scholars (e.g., Calor et al., 2022; Dekker & Elshout-Mohr, 2004) discerned two levels that influence teachers' support of group learning: content-specific and learning process. The content-specific support is usually aimed at helping groups succeed in the task, while learning process support assists individual students' learning progression. We modify this distinction to include content-specific and interpersonal factors as influences on group learning.

The content-specific factor includes long- or short-term students' characteristics. Long-term characteristics are usually traits or stable qualities of the students, such as gender, ethnicity, motivational level (Lei et al., 2010), learning style, and personality style (Chen & Kuo, 2019). For example, teachers use ability grouping to bring together students so that they learn throughout the year with others who are similar to them in personal characteristics (Dar, 1985; Slavin, 1987). Short-term content-specific characteristics usually indicate a student's performance in a small number of tasks or even a single task, which could be a result of manipulation or a genuine response. At the other end of the duration spectrum, as in the case of the present study, where students are grouped for a single activity, for example, think-pair-share (Kaddoura, 2013; Lyman, 1981) and peer instruction (Fagen et al., 2002; Zhang et al., 2017). Manipulated short-term characteristics often appear in group learning activities such as the JigSaw (Johnson & Johnson, 1989; 2002), where group members learn a specific topic with students external to their group, after which they return to their original group to teach their peers what they have learned, taking into account their shared learning goal. In situations of this type, the teacher understands what students have learned before the grouping. In many other cases, however, such as think-pair-share, the teacher has little control over what every student has learned.

The interpersonal criterion of grouping concerns the affective and cognitive relations between the students grouped. The affective relations between group members play a crucial role in group learning, as students often prefer to learn with their friends (e.g., Kutnick & Kington, 2005; Lou et al., 1996). Accordingly, friend grouping is considered productive for learning (Chapman et al., 2006; Chen & Kou, 2019) because friends are more inclined to share their thinking, voice disagreement (a central factor in promoting inquiry in learning), and take the time to listen to each other. Research on grouping also discerns between homogeneous or heterogeneous grouping (e.g., Maqtary et al., 2019; Webb, 2009).

Abdu et al. (2022) identified a trichotomous distinction to interpersonal considerations for grouping: *similarity, encompassment,* and *mutuality*. A homogeneous group usually comprises students who show *similarity* according to a chosen set of content-specific characteristics. A heterogeneous group is comprised of students who are considered dissimilar in a given set of content-specific characteristics. Heterogeneous groups contain one of the two *encompassing* or *mutual* relations between the group's participants' answers.

When students learn in a homogeneous group, it is more likely that they will progress at the same pace and be able to deal with tasks of similar difficulty. Ability grouping is an example of homogeneous grouping (e.g., Dar, 1985; Slavin, 1987). It has been used with Learning Analytics systems such as HMH READ 180© and Assessment to Instruction, in which teachers probe students' reading, comprehension, and vocabulary and receive a recommendation from a designated grouping module on how to form homogeneous groups and adapt the instruction to each learning group (Connor et al., 2013).

Typically, heterogeneous groups include an *encompassing* interaction where some students are meant to lead the interaction and guide their peers (Hoyles et al., 1991; Kontorovich et al., 2012). For example, in Wichmann et al. (2016), encompassing groups outperformed homogeneous ones, except when homogeneous groups comprised two strong students. Heterogeneity in grouping usually means focusing on levels of achievement, or what may be termed encompassing collaborative learning: high achievers who study with low achievers (e.g., Maqtary et al., 2019; Webb et al., 1997). Alternatively, it may mean focusing on other content-specific characteristics, such as socioeconomic status or age.

Another type of heterogeneous grouping may strive to achieve *mutuality,* by which each student has at least one unique content-specific characteristic that other members lack. In this way, students can combine their capabilities to tackle a particular learning task or teach each other what they know to expand their knowledge about a specific topic. This difference in thinking about a topic could be operationalised, as in the case of JigSaw, or discerned by the teacher/experimenter. Dialogic pedagogy emphasises individual differences as a source of learning through interaction (Abdu et al., 2021; Asterhan et al., 2020; Wegerif, 2011; Wegerif & Major, 2019). Accordingly, mutuality is the interpersonal condition that offers the most significant potential for learning from collaborative interaction because it fosters a dialogic gap that ignites bi-directional exchange and cognitive development (Abdu et al., 2022). Research on group learning consistently has shown that mutuality and mutual dependence among learners with equal status are crucial for successful collaborative learning (Cohen, 1994; Schwartz, 1999; Van den Bossche et al., 2011; Webb et al., 2014). Some researchers promote mutuality by deliberately eliciting or grouping students based on their differences at the outset of the interaction (Gutierrez-Santos et al., 2016; Schwarz et al., 2000). For instance, Schwarz et al. (2000) applied a Vygotskian approach to bug correction by pairing 10th graders with different misconceptions on a decimal fractions task. This approach was practical when the dyads employed productive group behaviors.

Using Learning Analytics for Content-specific Grouping

Content-specific grouping is the act of using data on students' knowledge of specific content and using these data to group them with other students. This tedious job requires collecting, analysing, and providing tailored information on the fly–an assignment almost impossible for teachers. Learning analytics can be instrumental in helping teachers monitor learners' thinking and learning (Raković et al., 2023; Stanja et al., 2023), provide insights into learners' cognitive processes and misconceptions (Yerushalmy et al., 2017), and support group learning (Abdu et al., 2022; Barron, 2003; Chen et al., 2019; Hoyles et al., 1991; Staples, 2008; Wise et al., 2021; Wise & Schwarz, 2017). Specifically, several learning analytics tools were developed to collect data about students' work to support informed formative assessment (D'angelo et al., 2015; Koedinger et al., 2010; Stacey & Wiliam, 2012; Yerushalmy et al., 2017). Such systems can support mathematics teachers' decision-making by probing students' perspectives (e.g., Segal et al., 2017; Yerushalmy et al., 2017). Such data may be used to make recommendations for teachers regarding students' grouping. Sinclair et al. (2011) used the term personal example space (PES) to signify one's contemporary, situated, idiosyncratic, and transient repertoire of available examples to think of, construct, and express a particular concept. Understanding how students access and generate mathematical examples can reveal invariant aspects of individuals' mathematical knowledge (Goldenberg & Mason, 2008). Accordingly, using learning analytics to probe for PES can give teachers a glance into the contemporary state of the concept a student chooses to express.

SYSTEM is an online learning analytics-based environment for automated formative assessment in mathematics. SYSTEM tasks are usually designed around interactive diagrams created with GeoGebra (Figure 1 & 2), a popular dynamic mathematics environment (Hohenwarter et al., 2009). Students solve the mathematical task and submit their answers, for example, as a finite solution to a question, as a set of examples supporting or refuting a mathematical claim, or as examples of a mathematical concept. SYSTEM collects the submitted answers and tags them based on predefined mathematical characteristics and on the correctness of the answer (Yerushalmy et al., 2017). The system then uses the data to report to students for further learning (Abdu et al., 2022) and the teacher to facilitate further instruction (Olsher et al., 2016). When students interact with feedback regarding their mathematical work, teachers can free up to support individuals' performances and interpersonal communication within group learning—such as grouping.

This paper is situated in a design-research project aiming to foster dialogic pedagogy by leveraging SYSTEM (Abdu et al., 2022; Olsher et al., 2025). Different task types could benefit from different grouping types (Olsher, 2022). However, mutuality grouping is the best fit for expanding individual students' PESs (Abdu et al., 2022). Olsher et al. (2025) showed that mathematics teachers attend to learning attributes of group learning considerably more than interpersonal attributes of group learning. We query the content-specific and interpersonal considerations that guide mathematics teachers when using data on students' learning to group students into study groups. Studies have shown that teachers often group students in ways that are not always congruent with pedagogical objectives (Kutnick & Kington, 2005). For example, Webb et al. (1997) showed that mathematics teachers often diverge from endorsed grouping strategies to allow friendship grouping and, at times, accommodate the demands of concrete tasks.

Implementing a SYSTEM to complement and support teachers' work requires us, as design researchers, to understand teachers' perspectives (Chapman, 2012) when focused on the students' mathematical work. Therefore, the main research question is:

> *What are mathematics teachers' considerations in grouping students based on data about students' answers to a rich mathematical task?*

Notably, we ask, (a) What are their considerations regarding the mathematical criterion for grouping? (b) What are their considerations regarding the interpersonal criterion for grouping? and (c) How do the mathematical and interpersonal criteria interact?

Methodology

Participants

Nine students enrolled in a Master's program at an Israeli University participated in this study. The program provides advanced research-oriented studies to both practicing and prospective teachers, along with their teacher certificate studies. The post-graduate students had a bachelor's degree in mathematics or mathematics education, had or were learning towards a high school teaching certificate in mathematics, and were familiar with learning analytics platforms (including SYSTEM). The levels of experience in teaching ranged from novice teachers to 20 years of teaching experience in elementary and secondary mathematics.

The Task

We used an example-eliciting task on the concept of the linear function (e.g., Moschkovich, 1996). Example-eliciting tasks are automated formative assessment tasks used to sample a student's PES by asking learners to create several examples of that concept (Yerushalmy, 2020). In the task, students were presented with a GeoGebra applet (Figure 1) and instructed:

> *Choose two red points using the "New points" button and use them to build a linear function whose graph passes through the points you have chosen. Submit three examples that are as different as possible.*

The presentations of the pairs of points in the applet were partially randomised, so the probability of the applet presenting some cases was higher than that of presenting others. For example, "one point is on the Y axis" appeared in 25% of cases, and "both points have the same Y value" appeared in 15% of cases. The expected example consists of a graph and an expression describing a linear function that passes through the two selected points.

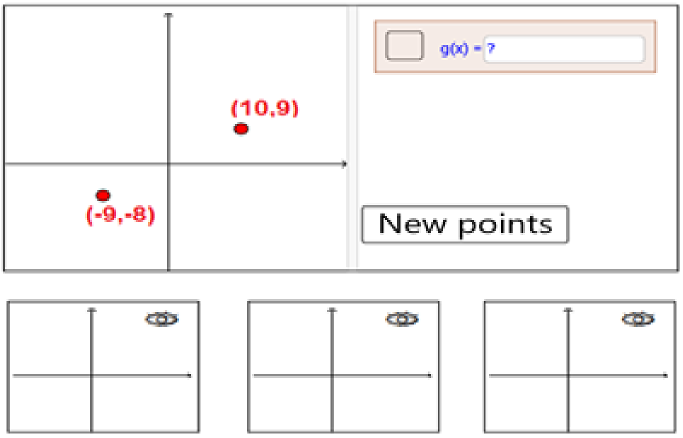

Figure 1. Example-eliciting task within the SYSTEM environment. The expected example consists of a graph and an expression describing a linear function that passes through the two selected points.

Several mathematical attributes can characterise the examples submitted for this task. SYSTEM provides learners and teachers with information regarding three attributes: orientation, construction method, and accuracy. (1) *Orientation* refers to the slope of the submitted example: ascending, descending, or constant linear function. (2) *Construction method* refers to the constraints the solver selects: one on the *x*- or the *y*-axis, the two have the same *x*-value, and the two have the same *y*-value. Each selection leads to a different solution path. For example, choosing a point on the *y*-axis supports immediate recognition of the b parameter in $y = a \cdot x + b$. However, when the two points coincide, the solver should create a single function out of infinite possibilities. (3) *Correctness* refers to the accuracy of the example submitted by the user: e.g., a graph that passes through the selected point(s). We identify two types of mistakes: one is when the line does not pass through one or both points; the other is when the graphic representation looks correct, but the algebraic expression does not represent a linear line passing through the required points (see Jones, 2000).

Procedure

A week before the experiment, we conducted a preparation for the experiment during a lesson in one of the program's courses. First, the course's students (20 students) solved the task individually and submitted their examples (i.e., personal example spaces) in SYSTEM. Next, they participated in a SYSTEM-based whole class discussion on the different examples provided by the participants—referring to the variety of orientations, methods of construction, and correctness in the participants' answers (Yerushalmy et al., 2017). We then prepared printouts of the 20 PES exemplars with fictitious student names produced in the preparation stage. We also created a form containing a table with three columns and eight rows. The headings of the rows were "Student A," "Student B," and "Why were they paired?" Hereafter, we refer to "students" by pseudonyms (e.g., Alex, Alina) and to teachers with lettering (e.g., Teacher-A, Teacher-G).

In the experiment, the nine participants first looked at the PES printouts and were given five minutes to think about possible ways of grouping those students yet wrote nothing. Next, the participants attended a 50-minute lesson on dialogism, grouping, and the role of mutuality in grouping. This intervention aimed to introduce the participants to the different types of interactions that different groupings might prompt. Participants then received the answer form and were given the following instructions.

> *Please group the students in pairs based on their answers to the mathematical task. The objective is to expand their PES, assuming the two students will later learn together. Explain your choice.*

The expectation was that teachers would focus on mathematical aspects of the students' answers, as the arbitrary nicknames freed them from considering personal relationships or other traits that would probably influence the teachers in grouping in the context of an actual classroom.

Data Analysis

The nine participants submitted 53 groupings (average 5.78; see Tables 1 & 2). We developed a coding scheme. First, the two authors created operational definitions for the items within the two criteria. Then, both coded the answers from two teachers and discussed their coding with an expert reviewer. Finally, the first author re-coded all the data according to the established coding scheme.

The *content-specific (mathematical) coding scheme* corresponded with the items represented in the task analysis above orientation, methods of construction, and correctness. A grouping by a teacher was coded as considering the *Construction method* characteristics when the teacher's explanation referred to specific points, such as: "the point chosen intersects the *y*- (or *x*-) axis." A grouping was coded as considering *Orientation* characteristics when the teacher's explanation contained one or more of the terms like "slope," "ascending," and "constant". A grouping was coded as considering *Correctness* when the teacher's explanation for the grouping contained terms such as "correct," "wrong answers," or their derivatives.

The *interpersonal coding scheme* included three relations between grouped students: similarity, encompassing, and mutuality. Grouping was coded as similarity when participants referred to the PES of each of the two paired students as similar, using words such as "same" or "alike" to describe the answers of the two students, together with an explanation of a proposal by the grouping teacher to develop the PES of both students in the same direction. Grouping was coded as *encompassing* when participants wrote *only* how one student could contribute to the other, and mutuality when participants wrote how each student could contribute to the other.

Findings

The Content-specific Criterion: Mathematical Considerations for Grouping

We recorded 66 instances of participants' content-specific considerations in the 53 groupings (Table 1). All groupings included at least one mathematical characteristic. In 12 cases, participants expressed two (11 cases) or all (1 case) content-specific considerations (see also Table 3). The most frequent consideration regarded function orientation (37 of 53). The *correctness* (15) and *construction* methods (13) appeared less frequently and not for all participants. Four participants considered all three aspects in their groupings.

Table 1.
Type and distribution of content-specific considerations for grouping by teachers.

Teacher	Number of groupings	Construction method	Orientation	Correctness
A	7	2	5	
B	5		5	1
C	7	3	4	1
D	8	2	4	3
E	6		6	1
F	4	2	3	1
G	7		4	6
H	4	2	3	
I	5	2	3	3
Total	53	13	37	16
Total content-specific considerations		66		

The most common consideration for the grouping was orientation, which included all explanations. For example, in the grouping of Hanna and Alex (Figure 2a & 2b), Teacher-A wrote:

> *Both created ascending functions, so I would like them to sit together to expand their example spaces and create more examples that are different from each other.*

Also, six teachers referred to *construction method* considerations in their grouping. For example, Teacher-F noted the work by Hanna, stating that "

> *Alex could teach Hanna about the intersection [of the function] with the axes, and Hanna could teach Alex about [choosing] a point on the x-axis [and] an example of a constant function.*

In Hanna's simplistic answer (Figure 2b), she chose only cases where points appear on the x-axis. For Teacher-F, she still can help expand Alex's PES. Seven teachers used *correctness* considerations when grouping. Grouping students based on the correctness of the functions submitted was mainly intended to ensure that students submit algebraically correct functions. Teacher-E addressed the correctness criterion in addition to orientation when he explained that:

> *Alex provided correct examples, although all were very similar (increasing functions), Hanna did not seem to be able to create the examples so that she could use help from Alex.*

By grouping Alex and Hanna, Teacher-E hoped Hanna would pay attention to creating correct functions. Note, however, the deficient insight by Teacher-E; the examples submitted by Alex were incorrect, as two of the three functions did not pass through the chosen points.

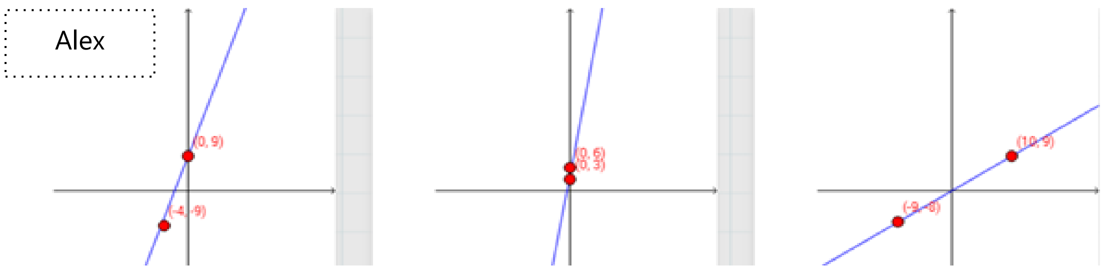

Figure 2a. Examples submitted by Alex.

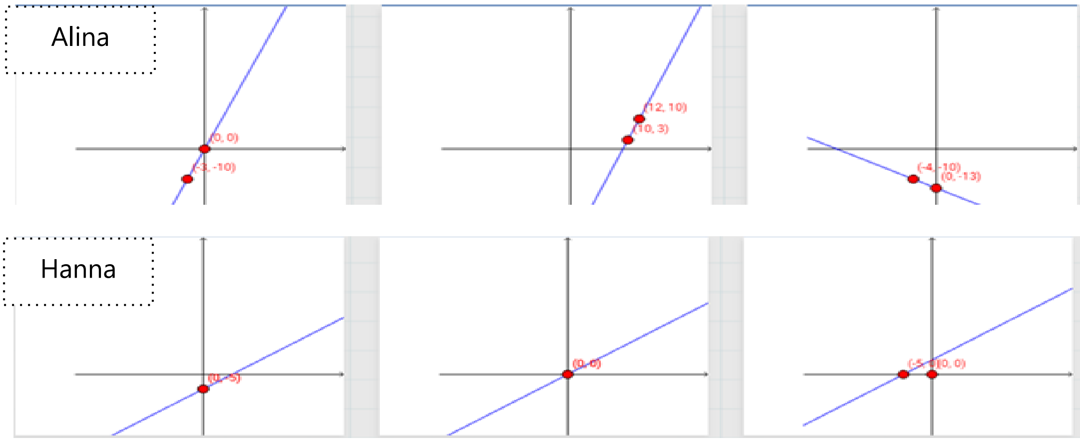

Figure 2b. Examples submitted by Alina and Hanna.

The Interpersonal Criterion of Grouping

Table 2 lists the groupings submitted by the teachers and the categories to which the groupings were assigned. The table shows that the most popular grouping type (33 out of 53) was encompassing-based. Encompassing was the one consideration used by all teachers. Mutuality and similarity were used only by seven and three teachers, respectively. Teachers A and C showed great versatility by grouping based on all three interpersonal considerations.

Table 2.
Type and Distribution of Interpersonal Considerations for Grouping by Teachers

Teacher	Number of groupings	Similarity	Encompassing	Mutuality
A	7	2	3	2
B	5		1	4
C	7	1	2	4
D	8		7	1
E	6		5	1
F	4		2	2
G	7		5	3
H	4		4	
I	5	1	4	
Total	53	3	33	17

Three teachers were grouped based on *similarity* considerations. For example, Teacher-A grouped Hanna and Alex, explaining,

> *Both of them created increasing functions, so I would like them to [...] expand their example spaces and create more examples that are different from each other.*

In similarity grouping, teachers often proposed a shared assignment for the pair as a group to expand the PES of both students in the same direction. Mainly by asking students to correct an example or expand their similar PESs concerning orientation or construction methods. All teachers grouped based on *encompassing* considerations—the most frequent interpersonal consideration. In these cases, teachers considered the submission of one student to be superior to that of another. For example, Teacher-E used correctness considerations, writing that:

> *Alex provided correct examples although all were very similar (increasing functions), Hanna did not seem to create all the examples so she could use help from Alex.*

Teacher-H used orientation considerations for encompassing grouping and writing,

> *Alex—the values of slopes are different; Hanna—the values of the slopes are similar.*

Seven teachers grouped based on *mutuality* considerations. In these cases, grouping was based on differences between students so that each may learn from the other. For example, Teacher-F grouped Alex and Alina (Figure 2a & 2b), using construction method considerations, writing:

> *Alex could teach Alina about the intersection [of the function] with the axes, and Alina could teach Alex about [choosing] a point on the x-axis [and creating] an example of a constant function.*

Same Grouping, Different Interpersonal Considerations

Nine pairs appeared in more than one teacher's grouping. Twice, in seven cases, these considerations were coded differently. Five groupings appeared twice, thrice by three teachers and once by four teachers. For example, Hanna and Alex (Figure 2a & 2b) were grouped by Teacher-A based on similarity considerations and by Teacher-E and Teacher-H based on encompassing considerations. In his explanation, Teacher-A mentioned only orientation characteristics (see the example above). Teacher-H also used orientation characteristics but acknowledged that the slopes presented by Hanna were all similar, whereas Alina's slopes varied. Teacher-E, however, used two criteria: orientation ("increasing functions") and correctness ("Alex provided correct examples"). None of the teachers in these examples acknowledged the construction method in Hanna's examples. Other teachers (e.g., Teacher-D) could identify that in two cases, Hanna chose two points that merged. All three example functions for both students were ascending. However, Hanna used two methods of construction: selecting a point on the *x*-axis (0, -5) and selecting two points that overlap (0, 0); notably, point (0, 0) could also be counted as the first construction method. Hanna also created two correct functions out of three, and Alex created none.

Combining Interpersonal and Content-specific

We collapsed the content-specific and interpersonal criteria into one table to consider overall teachers' grouping strategies based on both criteria (Table 3). We identified two phenomena that shed light on teachers' grouping strategies: leading grouping strategies and the distinction between within- and between content-specific considerations for grouping.

Table 3.
Content-specific (rows) vs. Interpersonal (columns) Considerations.

	Similarity	Encompassing	Mutuality	Total
Heuristic	2 (4%)	3 (6%)	1 (2%)	6 (11%)
Orientation	2 (4%)	14 (26%)	11 (21%)	27 (51%)
Correctness		7 (13%)	1 (2%)	8 (15%)
Heuristic & Orientation		1 (2%)	3 (6%)	4 (8%)
Orientation & Correctness		5 (9%)		5 (9%)
Heuristic & Correctness		2 (4%)		2 (4%)
Heuristic, Orientation, & Correctness		1 (2%)		1 (2%)
Total	4 (~8%)	32 (~61%)	16 (~31%)	53

Leading grouping strategies

Three content-specific and interpersonal combined grouping strategies were most frequent: encompassing-orientation, mutual-orientation, and encompassing-correctness (Table 3). Encompassing-orientation grouping strategies were used when the submissions by one student included functions in more orientations than those of the other student. For example, Teacher-H grouped Hanna and Alex, explaining that:

> *Alex—the slopes are different; Hanna—the slopes are same.*

Mutual-orientation grouping strategies were used when each student submitted at least one function in an orientation neglected by the other. For example, Teacher B explained grouping Alina (Figure 2b) with Sasha (Figure 3):

> *Alina manages to create constant functions; Sasha manages to create decreasing-increasing functions. [They] can learn from each other.*

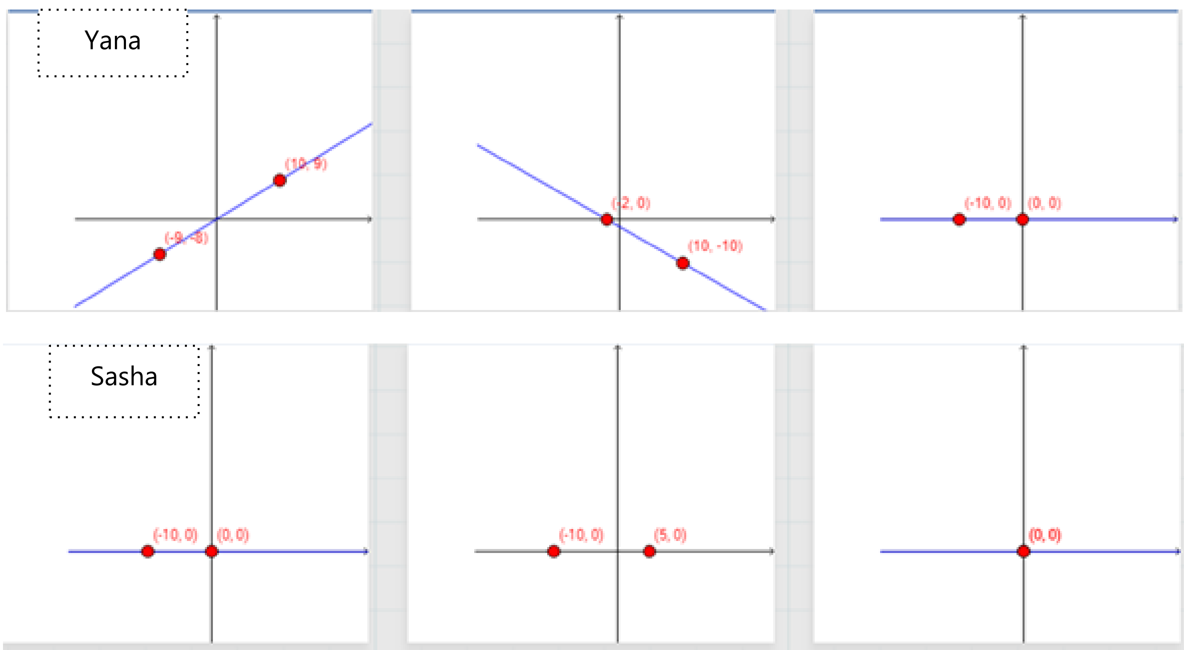

Figure 3. The PESs of Yana and Sasha.

Encompassing-correctness grouping strategies were used when the functions submitted by one student were correct more often than those of the other student. For example, Sasha was grouped with three different students by three different teachers (C, D, and F). In all three cases, Sasha was grouped with a student who submitted functions that did not pass through the points selected.

Within- and between-content-specific Considerations for Grouping

Teachers grouped students using within- and between-content-specific considerations. Regarding within-content-specific considerations, the teacher explained only one content-specific consideration for the grouping, like the ones described above. The mutual-orientation grouping strategy is an example of this approach. By contrast, the teacher explained more than one content-specific consideration for the grouping in the case of between-content-specific considerations. For example, Teacher-G explained that she grouped Yana (Figure 3) and Hanna (Figure 2b) because

> *Yana solved correctly and with variance* [between answers]. *Hanna did not solve the exercises correctly, so the first may teach the second.*

Note that an automated grouping module would have analysed this case otherwise. Teacher-G performed encompassing grouping based on correctness and orientation considerations. However, based on the construction method consideration, it was noted that Yana could learn from Hanna, as the latter resorted to construction methods such as choosing points that overlap and a point on the *y*-axis.

Discussion and Conclusion

We aimed to understand mathematics teachers' considerations in using data on students' performance in mathematically rich tasks. What are mathematics teachers' considerations in grouping students based on data about students' answers to a rich mathematical task? In doing so, we created a dual framework combining content-specific and interpersonal considerations for grouping.

In general, teachers consider interpersonal cognitive characteristics mostly in long-term grouping. When students are to be grouped for shorter time frames, such as a single task, teachers may not be able to take the time to collect data on all students concerning predefined characteristics and perform optimal grouping (Lou et al., 1996; Maqtary et al., 2019; Slavin, 1987). For teachers who do not use automated tools to collect data on students' work, grouping for short time frames, as suggested in the present paper, may be perceived as inefficient and not worth investing the extra effort. Performing a meticulous analysis of students' answers to a given task to group students for a single additional task may seem too laborious. We argue that automated grouping may be a productive way of approaching this problem through design. In the *content-specific criterion*, teachers chose orientation considerations more than construction method and correctness. Mathematics teachers are usually trained to identify incorrect answers and attempt to correct them (Stacey et al., 2009). In turn, the pedagogical approach underlying probing for students' PES assumes teachers can consider multiple characteristics in students' examples to a mathematical concept (Yerushalmy et al., 2017). In our experiment, 45 out of 55 examples submitted were correct, so teachers may have looked past the correctness characteristic. The finding that orientation was central for grouping, followed by construction method, contributes to research on function recognition by experts (e.g., Kop et al., 2017), suggesting that the orientation of a linear function is more evident and relevant to teachers than the characteristics referred to as construction method.

For the *interpersonal criterion*, we advocated mutuality consideration with a request before the grouping to "*group students so that all the students would be able to teach their partners*." Mutuality grouping is consistent with central ideas in research on group learning that advocate the importance of reciprocity between students (Cohen, 1994; Schwarz et al., 2015; Webb, 2009; Webb et al., 2014) of differences between students (Wegerif, 2011), and of the agency of group members (Schwartz, 1999) for supporting productive group learning. In this study, however, teachers grouped students primarily based on encompassing considerations, occasionally based on mutuality considerations, and only rarely based on similarity considerations. That is, the mutuality-happy request was only partially effective. In many cases, teachers did not acknowledge differences between grouped students, discerning encompassing relations as similar and mutual relations as encompassing or similar. We propose two explanations for this phenomenon. First, teachers' epistemic practices may be congruent with most research on grouping that champions the homogeneous-heterogeneous dichotomy (e.g., Dar, 1985; Lou et al., 1996; Maqtary et al., 2019; Slavin, 1987) and perhaps influenced by this research. This dichotomy may affect their epistemic practices in ways that prevent teachers from accepting the idea of mutuality in grouping or because they perceive the tasks' goals will be best met with similarity or encompassing grouping (Abdu et al., 2022). A second explanation for the lack of complexity in grouping may be teachers' ability to conduct grouping based on complex considerations (e.g., mutual grouping, between-tasks considerations) in short time frames. Such practice requires knowledge of the mathematical task, discerning content-specific characteristics, prioritising them according to the teacher's objectives, assessing all the students in the class based on these characteristics, and deciding which students should learn together. Processing large amounts of data quickly may overwhelm teachers, undermining their effort to conduct well-considered groupings.

At the intersection between the content-specific and interpersonal criteria, we identified three grouping strategies that were more popular than others, namely, encompassing-orientation, mutual-orientation, and encompassing-correctness. As noted above, orientation was the focus of many groupings. Teachers tried to group students in ways that would diversify orientations. In other cases, they grouped students who answered incorrectly with those who answered correctly to eliminate incorrect solutions. In rare instances, teachers performed complex groupings. In 12 of the 53 groupings, teachers considered two or three content-specific characteristics, and in 9 of those 12 groupings, the two content-specific characteristics were both used to explain the superiority of one student over the other (encompassing); alternatively, one characteristic served as a baseline to indicate when answers were similar, and the other characteristic to indicate how one student could contribute to the other.

Teachers focused on different characteristics in their grouping—in nine cases, two or more teachers grouped the same two students using different content-specific and interpersonal considerations. This attests to the variability in teachers' possibilities and considerations concerning the desired outcomes. Webb et al. (1997) showed that teachers do not always follow recommendations for grouping because they consider issues that are not necessarily addressed in controlled studies, such as familiarity between students and maintaining low diversity among them. In support, this study shows the same phenomena in a "cleaner" circumstance that focuses on content-specific characteristics and omits personality considerations.

Grouping students based on their performance in each task, using fine-grained considerations such as those elaborated above, is a meticulous and tedious job. Automated tools such as SYSTEM may offload this work from teachers (see also Gutierrez-Santos et al., 2016) by modelling content-specific and interpersonal considerations. Data collected on students' answers to designated tasks, like the one presented above, can be used to group students using complex considerations.

Further studies should inquire to what extent automated grouping modules promote productive group and individual learning. Similarity and encompassing grouping strategies may be practical in other contexts where the purpose is not the enrichment of a PES. For example, similarity could be practical in settings that need both students to solve a problem together as a homogeneous group, and encompassing may be effective in situations where students have mastered a practice and others have not yet.

In our effort to automate the process of grouping, we sought to understand how teachers used SYSTEM-based data to propose ways of grouping. The present study has contributed to creating effective automated grouping modules using automated learning analytics systems. Understanding this process can help teachers foster effective student-centered mathematics tasks for learning mathematical content.

Corresponding author

Rotem Abdu
David Yellin Academic College of Education
rotem.abdu@gmail.com

Ethical approval

Ethical approval for the research was granted by an author's institution, and informed consent was given by all participants for their data to be published. Ethical approval was granted by the Israeli Ministry of Education (11023#).

Competing interests

The authors declare there are no competing interests.

References

Abdu, R., Schwarz, B., & Mavrikis, M. (2015). Whole-class scaffolding for learning to solve mathematics problems together in a computer-supported environment. *ZDM Mathematics Education*, *47*(7), 1163–1178. https://doi.org/10.1007/s11858-015-0719-y

Abdu, R., van Helden, G., Alberto, R., & Bakker, A. (2021). Multimodal dialogue in small-group mathematics learning. *Learning, Culture and Social Interaction*, *29*, Article 100491. https://doi.org/10.1016/j.lcsi.2021.100491

Abdu, R., Olsher, S., & Yerushalmy, M. (2022). Pedagogical considerations for designing automated grouping systems: The case of the parabola. *Digital Experiences in Mathematics Education, 8*(1), 99–124. https://doi.org/10.1007/s40751-021-00095-7

Asterhan, C. S., Howe, C., Lefstein, A., Matusov, E., & Reznitskaya, A. (2020). Controversies and consensus in research on dialogic teaching and learning. *Dialogic Pedagogy*, *8*. https://doi.org/10.5195/dpj.2020.312

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development*, *2*(3), 202–221. https://doi.org/10.1111/j.1467-9507.1993.tb00014.x

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12*(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Calor, S. M., Dekker, R., van Drie, J. P., & Volman, M. L. (2022). Scaffolding small groups at the group level: Improving the scaffolding behavior of mathematics teachers during mathematical discussions. *Journal of the Learning Sciences*, *31*(3), 369-407. https://doi.org/10.1080/10508406.2021.2024834

Chapman, O. (2012). Challenges in mathematics teacher education. *Journal of Mathematics Teacher Education*, *15*(4), 263–270. https://doi.org/10.1007/s10857-012-9223-2

Chapman, K. J., Meuter, M., Toy, D., & Wright, L. (2006). Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes. *Journal of Management Education*, *30*(4), 557–569. https://doi.org/10.1177/1052562905284872

Chen, C. M., & Kuo, C. H. (2019). An optimized group formation scheme to promote collaborative problem-based learning. *Computers & Education, 133*, 94–115. https://doi.org/10.1016/j.compedu.2019.01.011

Chen, L., Yoshimatsu, N., Goda, Y., Okubo, F., Taniguchi, Y., Oi, M., Konomi, S., Shimada, A., Ogata, H., & Yamada, M. (2019). Direction of collaborative problem-solving-based STEM learning by learning analytics approach. *Research and Practice in Technology Enhanced Learning*, *14*(1), Article 24. https://doi.org/10.1186/s41039-019-0119-y

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1–35. https://doi.org/10.3102/00346543064001001

Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, *24*(8), 1408–1419. https://doi.org/10.1177/0956797612472204

D'angelo, C. M., Roschelle, J., & Bratt, H. (2015). Using students' speech to characterize group collaboration quality. In proceedings of the *11th International Conference on Computer Supported Collaborative Learning: Exploring the Material Conditions of Learning, CSCL 2015* (pp. 819–820). International Society of the Learning Sciences (ISLS).

Dar, Y. (1985). Teachers' attitudes toward ability grouping: Educational considerations and social and organizational influences. *Interchange*, *16*(2), 17–38. https://doi.org/10.1007/BF01807206

Dekker, R., & Elshout-Mohr, M. (2004). Teacher interventions aimed at mathematical level raising during collaborative learning. *Educational Studies in Mathematics*, *56*(1), 39–65.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.

Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher*, *40*(4), 206–209. https://doi.org/10.1119/1.1474140

Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics*, *69*(2), 183–194. https://doi.org/10.1007/s10649-008-9143-3

Gutierrez-Santos, S., Mavrikis, M., Geraniou, E., & Poulovassilis, A. (2016). Similarity-based grouping to support teachers on collaborative activities in an exploratory mathematical microworld. *IEEE Transactions on Emerging Topics in Computing*, *5*(1), 56–68. https://doi.org/10.1109/TETC.2016.2533318

Hopfenbeck, T. N., Zhang, Z., Sun, S. Z., Robertson, P., & McGrane, J. A. (2023). Challenges and opportunities for classroom-based formative assessment and AI: A perspective article. *Frontiers in Education, 8*, Article 1270700. https://doi.org/10.3389/feduc.2023.1270700

Hohenwarter, J., Hohenwarter, M., & Lavicza, Z. (2009). Introducing dynamic mathematics software to secondary school teachers: The case of GeoGebra. *Journal of Computers in Mathematics and Science Teaching*, *28*(2), 135–146. https://www.learntechlib.org/p/30304/

Hoyles, C., Healy, L., & Sutherland, R. (1991). Patterns of discussion between pupil pairs in computer and non-computer environments. *Journal of Computer Assisted Learning*, *7*(4), 210–228. https://doi.org/10.1111/j.1365-2729.1991.tb00253.x

Johnson, D. W. & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction.

Johnson, D. W., & Johnson, R. T. (2002). Learning together and alone: Overview and meta-analysis. *Asia Pacific Journal of Education*, *22*(1), 95–105. https://doi.org/10.1080/0218879020220110

Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics*, *44*(1-2), 55–85. https://doi.org/10.1023/A:1012789201736

Kaddoura, M. (2013). Think pair share: A teaching learning strategy to enhance students' critical thinking. *Educational Research Quarterly*, *36*(4), 3–24.

Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, *43*(4), 489–510. https://doi.org/10.2190/EC.43.4.d

Kontorovich, I., Koichu, B., Leikin, R., & Berman, A. (2012). An exploratory framework for handling the complexity of mathematical problem posing in small groups. *The Journal of Mathematical Behavior*, *31*(1), 149–161. https://doi.org/10.1016/j.jmathb.2011.11.002

Kop, P. M., Janssen, F. J., Drijvers, P. H., & van Driel, J. H. (2017). Graphing formulas: Unraveling experts' recognition processes. *The Journal of Mathematical Behavior*, *45*, 167–182. https://doi.org/10.1016/j.jmathb.2017.01.002

Kutnick, P., & Kington, A. (2005). Children's friendships and learning in school: Cognitive enhancement through social interaction? *British Journal of Educational Psychology*, *75*(4), 521–538. https://doi.org/10.1348/000709904X24591

Lei, S. A., Kuestermeyer B. N., & Westmeyer, K. A. (2010). Group composition affecting student interaction and achievement: Instructors' perspectives. *Journal of Instructional Psychology*, *37*(4), 317–325.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*(4), 423–458. https://doi.org/10.3102/00346543066004423

Lyman, F. (1981). The responsive classroom discussion. In A. S. Anderson (Ed.), *Mainstreaming Digest* (pp. 109–113). University of Maryland College of Education.

Maqtary, N., Mohsen, A., & Bechkoum, K. (2019). Group formation techniques in computer-supported collaborative learning: A systematic literature review. *Technology, Knowledge and Learning*, 24(2), 169–22.

Moschkovich, J. N. (1996). Moving up and getting steeper: Negotiating shared descriptions of linear graphs. *The Journal of the Learning Sciences*, *5*(3), 239–277. https://doi.org/10.20935/AcadMatSci6212

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.

Olsher, S. (2022). Te (a) Ching to collaborate: Automatic assessment-based grouping recommendations and implications for teaching. In *Proceedings of the 15th International Conference on Technology in Mathematics Teaching* (p. 214). Aarhus University.

Olsher, S., Abdu, R., & Shalata, M. (2025). The relationships between student content-specific grouping and teachers-students' interactions during online collaborative mathematical task solving. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-024-10382-w

Raković, M., Gašević, D., Hassan, S. U., Ruipérez Valiente, J. A., Aljohani, N., & Milligan, S. (2023). Learning analytics and assessment: Emerging research trends, promises and future opportunities. *British Journal of Educational Technology, 54*(1). 10–18.

Schwarz, B. B., De Groot, R., Mavrikis, M., & Dragon, T. (2015). Learning to learn together with CSCL tools. *International Journal of Computer-Supported Collaborative Learning*, *10*(3), 239–271. https://doi.org/10.1007/s11412-015-9216-0

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right...If they argue together! *Cognition and Instruction*, *18*(4), 461–494. https://doi.org/10.1207/S1532690XCI1804_2

Schwartz, D. L. (1999). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (p. 197–218). Elsevier.

Sinclair, N., Watson, A., Zazkis, R., & Mason, J. (2011). The structuring of personal example spaces. *The Journal of Mathematical Behavior*, *30*(4), 291–303. https://doi.org/10.1016/j.jmathb.2011.04.001

Segal, A., Hindi, S., Prusak, N., Swidan, O., Livni, A., Palatnic, A., & Schwarz, B. (2017, June). Keeping the teacher in the loop: Technologies for monitoring group learning in real-time. In *International Conference on Artificial Intelligence in Education* (pp. 64–76). Springer.

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research, 50*(2), 315–342. https://doi.org/10.3102/00346543050002315

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, *57*(3), 293–336. https://doi.org/10.3102/00346543057003293

Stacey, K., Price, B., Steinle, V., Chick, H., & Gvozdenko, E. (September 28–October 1, 2009). SMART assessment for learning. Paper presented at the *Conference of the International Society for Design and Development in Education,* Cairns, Australia.

Stacey, K., & Wiliam, D. (2012). Technology and assessment in mathematics. In M. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & F. Leung, F. (Eds.), *Third international handbook of mathematics education* (pp. 721–751). Springer. https://doi.org/10.1007/978-1-4614-4684-2_23

Stanja, J., Gritz, W., Krugel, J., Hoppe, A., & Dannemann, S. (2023). Formative assessment strategies for students' conceptions—The potential of learning analytics. *British Journal of Educational Technology*, *54*(1), 58–75. https://doi.org/10.1111/bjet.13288

Staples, M. E. (2008). Promoting student collaboration in a detracked, heterogeneous secondary mathematics classroom. *Journal of Mathematics Teacher Education*, *11*(5), 349–371. https://doi.org/10.1007/s10857-008-9078-8

Van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team learning: Building shared mental models. *Instructional Science*, *39*(3), 283–301. https://doi.org/10.1007/s11251-010-9128-3

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *The British Journal of Educational Psychology*, *79* (1), 1–28.

Webb, N. M., Baxter, G. P., & Thompson, L. (1997). Teachers' grouping practices in fifth-grade science classrooms. *The Elementary School Journal*, *98*(2), 91–113. https://doi.org/10.1086/461886

Webb, N. M., Franke, M. L., Ing, M., Wong, J., Fernandez, C. H., Shin, N., & Turrou, A. C. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research*, *63*, 79–93. https://doi.org/10.1016/j.ijer.2013.02.001

Wegerif, R. (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity*, *6*(3), 179–190. https://doi.org/10.1016/j.tsc.2011.08.002

Wegerif, R., & Major, L. (2019). Buber, educational technology, and the expansion of dialogic space. *AI & Society*, *34*, 109–119.

Wichmann, A., Hecking, T., Elson, M., Christmann, N., Herrmann, T., & Hoppe, H. U. (2016, August). Group formation for small-group-learning: Are heterogeneous groups more productive? In *Proceedings of the 12th International Symposium on Open Collaboration*, Article 14. https://doi.org/10.1145/2957792.2965662

Wise, A. F., Knight, S., & Ochoa, X. (2021). What makes learning analytics research matter. *Journal of Learning Analytics*, *8*(3), 1–9. https://doi.org/10.18608/jla.2021.7647

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, *12*(4), 423–467. https://doi.org/10.1007/s11412-017-9267-5

Yerushalmy, M., Nagari-Haddif, G., & Olsher, S. (2017). Design of tasks for online assessment that supports understanding of students' conceptions. *ZDM Mathematics Education*, *49*, 701–716. https://doi.org/10.1007/s11858-017-0871-7

Yerushalmy, M. (2020). Seeing the entire picture with STEP: An example-eliciting approach to online formative assessment. In B. Barzel & F. Schacht (Eds.), *Proceedings of the 14th International Conference on Technology in Mathematics Teaching,* Essen, Germany, July 2019. https://nbn-resolving.org/urn:nbn:de:hbz:464-20191119-153831-9

Zhang, P., Ding, L., & Mazur, E. (2017). Peer instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs. *Physical Review Physics Education Research, 13*(1), Article 010104. https://doi.org/10.1103/PhysRevPhysEducRes.13.010104