Pedagogical Considerations for Designing Automated Grouping Systems: the Case of the Parabola

Rotem Abdu¹ · Shai Olsher¹ · Michal Yerushalmy¹

Accepted: 26 August 2021 / Published online: 22 September 2021

© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021, corrected publication 2021

Abstract

This article queries how learning analytics systems can support content-specific group formation to develop students' thinking about a specific mathematical concept. Automated group formation requires identifying personal characteristics, designing tasks to probe students' perceptions, and grouping them to increase individual learning chances. Designers of automated group formation recommendation modules (GFRMs) rarely consider content-specific objectives. We draw on theories on conceptual learning in mathematics and dialogic thinking to emphasize the role of a dialogic gap between students' voices to enhance individual learning. In an experiment, fifty 8th and 9th grade students solved three mathematical tasks in a pre-intervention-post-set-up: individually, then in dyads, and then individually again. We used a learning analytics system to collect fine-grained content-specific data on students' responses based on four pre-defined aspects of the parabola concept. We compared students' answers with those of their peers in order to identify interpersonal relations. The experiment results indicate that students' thinking about the parabola concept was the most successfully developed when every group member had a different perception of this concept. We illustrate the learning trajectories of four students and elaborate on the learning sequence of one of these students in particular. This article suggests that the centrality of a dialogic gap in developing personal learning is probably content independent. We thus call for software engineers to think about GFRMs that can support content-specific learning and instruction.

Keywords Group formation · Learning analytics · Personal example space · Example-eliciting tasks · Dialogic pedagogy

The composition of a group—the learners and how they are related—is an essential factor influencing individual learning effectiveness (Lou et al., 1996). Accordingly,

✉ Rotem Abdu
rotem_abdu@yahoo.com

¹ University of Haifa, Haifa, Israel

group formation is a process in which individual learners come to learn together—peers are self-selected by group members or elicited by an external entity, such as a teacher and/or technology (Zheng et al., 2018). A thoughtful group formation by an external entity requires tapping on *personal* and *interpersonal* considerations to increase the possibility of the group's composition leading to positive learning outcomes from a task or activity or over longer learning periods (Lou et al., 1996; Maqtary et al., 2019; Pearlstein, 2021).

Manual group formation may be very time-consuming, because it requires collecting and analyzing large amounts of data about students and then considering this data to optimize learning chances (Abdu et al., 2019; Erkens et al., 2016). This challenge led computer scientists and learning scientists to engineer group formation recommendation modules (GFRMs). These are algorithms that can provide instructors and students with recommendations for which learners should learn together (Amara et al., 2016; Borges et al., 2018; Holmberg, 2019; Isotani et al., 2009; Konert et al., 2014; Meyer, 2009; Moreno et al., 2012; Wang et al., 2007; Zheng et al., 2018).

GFRMs serve a great need in solving an old pedagogical problem—providing recommendations about which students should learn together. However, there is a scarce discussion in the literature about the pedagogical opportunities and implications of using GFRMs to support content-specific learning. Only a handful of studies discuss content-specific pedagogical considerations to such systems' design, let alone rooting this design in learning theories (Borges et al., 2018; Maqtary et al., 2019).

This article addresses this lacuna by focusing on the pedagogical considerations to design a content-specific GFRM. We identify the type of *tasks* that can probe and develop students' thinking about the concept of parabola, characterize *personal data* (that could be) collected in relation to these tasks to reflect students' thinking, and study how certain *interpersonal set-ups* influence the students' learning.

Theoretical Background

Content-specific GFRMs require probing for students' current thinking about that content before grouping. The *personal considerations* for grouping define the data to be gathered about individuals in light of a learning goal. The type of data scientists who develop GFRMs can gather how learners' traits govern these personal considerations: for example, overall learning skills, gender and personality types (Zheng & Pinkwart, 2014), domain-specific skills such as overall programming (Zheng et al., 2018), and social skills such as leadership and social interaction (Moreno et al., 2012).

As mathematics educators, we are often interested in fostering students' thinking about a particular concept. Therefore, content-specific GFRMs should enable a probing of students' thinking about that concept to “decide” with whom they should learn. However, what does it mean to think about a particular concept? Studies on conceptual learning in STEM education often use terms such as *conception* (Posner

et al., 1982), *concept image* (Vinner, 1983), and *personal example space* (Sinclair et al., 2011) to understand learners' thinking about a specific concept.

In particular, a *personal example space* (PES) can be seen as a student's current repertoire of available examples to think of, construct, and express a concept. Adopting this perspective means gaining information about how mathematical examples are accessed and generated so that they can reveal invariant aspects of individuals' mathematical knowledge (Goldenberg & Mason, 2008; Sinclair et al., 2011). Such a space is situated, idiosyncratic, and transient in and through learners' interaction with the environment. Probing for one resembles taking a snapshot of the momentary state of the concept as manifested by the student. A sequence in which such spaces change over time can model learning relating to that target concept.

Learning analytics can probe students' PESs about mathematical concepts (e.g., Olsher et al., 2016) and it requires eliciting students' thinking in the context of a particular concept and collecting data about their answers. In our laboratory at the University of Haifa, we developed an automated formative assessment system we termed *Seeing the Entire Picture* (STEP) (Olsher et al., 2016). The type of task we use to probe for students' PESs is called an *example-eliciting task*—open tasks in which students are asked to generate several examples of a mathematical concept under a closed set of constraints (Yerushalmy et al., 2017; Yerushalmy & Olsher, 2020).

Teachers design such tasks to probe for such spaces: For instance, students can be asked to choose three pairs of points and build three quadratic functions that pass through these pairs of points (see Figs. 2, 3, 4, and 5). The example-eliciting tasks in STEP are all based on interactive platforms (*GeoGebra*; see Hohenwarter et al., 2009) in order to gather data automatically about students' PESs. What kind of data can be collected on students' spaces with example-eliciting tasks?

In the context of a specific task, examples submitted by students in STEP can be analyzed along with a pre-defined set of mathematical aspects with which a mathematical concept may vary (Olsher et al., 2016). STEP can provide teachers and students with data about the characteristics of every submitted example along with these pre-defined mathematical aspects. For example, mathematics educators can analyze the mathematical concept of parabola along the mathematical aspect of "roots" (the number of intersections with the x -axis), and this aspect can have one of three possible values (0, 1, and 2). Identifying the mathematical characteristics of several examples submitted by a single student can shed light on their thinking about that concept. It can highlight recurring patterns in the examples (e.g., all their parabolas had one root) to discern the "comfort zones" of the student. Teachers can also identify variance among the submitted examples to discern the "boundaries" of the PES as manifested in the student's examples. The space notion does not signify what the student knows or understands—it represents a materialistic process that focuses on the examples a student chooses to submit.

Learning analytics systems such as STEP can support content-specific automated formative assessment by collecting data about PESs in the context of a specific concept. The ability to extract such data can enable content-specific automated group formation of many participants and data entries. They can also track personal learning regarding PES changes due to interaction with materials and peers.

Interpersonal Considerations for Group Formation

Collaborative learning is considered productive for individual learning, although some group learning situations are more productive than others (Barron, 2003; Cohen, 1994; Dillenbourg, 1999; Johnson & Johnson, 2002). Much work has been done to understand how different factors influence the effectiveness of collaborative learning. These factors include students' participation and social behavior when they learn together (e.g., Abdu et al., 2019; Barron, 2003; Dillenbourg, 1999; Kontorovich et al., 2012; Webb et al., 2014), effects of teacher support (Dekker & Elshout-Mohr, 2004; van Leeuwen & Janssen, 2019; Webb, 2009), and, specifically for the current article, the effects of different group formations on the effectiveness of learning (e.g., Lou et al., 1996; Pearlstein, 2021). Educators can use learning analytics systems to support collaborative learning (Wise & Schwarz, 2017): For example, by presenting data about the indicators of collaborative learning situations (D'Angelo et al., 2015; Schwarz et al., 2018; Wise & Schwarz, 2017), supporting teachers' decision-making in regard to collaborative learning (Martinez-Maldonado, 2019; van Leeuwen, 2015), supporting co-reflection on the collaborative learning process (Schwarz et al., 2015) and, specifically to this current article, group formation (see Borges et al., 2018; Maqtary et al., 2019).

The *interpersonal considerations* for grouping are rules governing how relations between students (i.e., *interpersonal set-ups*) inform who studies with whom. Taking together, literature on collaborative learning and dialogue (e.g., Cohen, 1994; Dillenbourg, 1999; Schwarz et al., 2000; Tsovaltzi et al., 2019; Wegerif, 2011), group formation (e.g., Lou et al., 1996; Slavin, 1987), and automated group formation (e.g., Amara et al., 2016; Maqtary et al., 2019), we identify three mutually exclusive interpersonal set-ups in collaborative learning: *similarity*, *encompassing*, and *mutuality* (see Fig. 1).

Fig. 1 A schematic view of three possible interpersonal set-ups student A may encounter. This example presents three possible interpersonal set-ups by student A with, in turn, students B, C, and D. Every column represents a PES. A learning analytics system can identify every such space along some aspects (four in this example). The value of every aspect can be one of several characteristics and a student may (filled rectangles) or may not (white rectangles) manifest characteristics associated with these aspects

Similarity (also known as *homogeneity*, Lou et al., 1996) represents set-ups in which all the students in a group were at a close developmental level and performed the same when data was collected. Groups of students with similar developmental levels allow teachers to treat them as a coherent pedagogical unit (e.g., Berland et al., 2015; Connor et al., 2013).

Encompassing (mostly known as *heterogeneity*, Lou et al., 1996) represents interpersonal set-ups in which (along with a pre-defined set of aspects), before the group formation, one or more students in the group performed better than their peers. In the context of PES, this means exemplifying one for a certain concept that encompasses his peer's comparable space. Encompassing set-ups are applied mostly when individual students have mastered practices that other group members have not yet (Lou et al., 1996). This process resembles a Vygotskian zone of proximal development, in which a knowledgeable person supports student development (Vygotsky, 1978).

Mutuality (for example, jigsaw set-ups; Cohen, 1994) represents interpersonal set-ups in which every student in the group manifests singular characteristics relevant to the task that was not produced by any of the other group members when the data was collected (Erkens et al., 2016). Mutuality interpersonal set-ups correspond with ideas from dialogic pedagogy, as illustrated in the next sub-section.

Operationalizing Dialogic Pedagogy for Content-Specific GFRM

Dialogic theories and their educational derivatives predict how interpersonal set-ups may influence content-specific collaborative learning. Inspired by certain philosophers (e.g., Bakhtin, 1984; Buber, 1923), dialogic pedagogy research focuses on the relationships between interpersonal interactions and individual learning. Bakhtin-inspired approaches (henceforth, dialogic pedagogy) often exemplify learning as a sequence of changes in a learner's perspective that emerges in and through interaction with other people and/or cultural agents (e.g., Asterhan et al., 2020; Teo, 2019; Wegerif, 2011).

An idea central to dialogic pedagogy is the voice, “the intentions [...] individual speakers present in each utterance” (Barwell, 2016, p. 335). A voice is situated—it is always about “something” (e.g., an idea, concept, or opinion) (Wegerif, 2011); it is contingent due to immanent continuous interaction with other voices (Bakhtin, 1984); it is subjective and multimodal (Abdu et al., 2021); it can be changed in interaction with cultural agents manifested in people and technology (Wegerif & Major, 2019). A *dialogue* happens when and where two or more *different* voices interact. In a classroom, a teacher, peer, or technology can express different voices. When students learn in collaboration within classroom settings, teachers may need to suppress their voices, allowing students to interact (Webb, 2009). Learning becomes external to the teacher—the classroom contains not a single objective curricular truth held together by the teacher's voice but a plurality of consciences, each with its own voice. Every student can potentially make a singular contribution to the interaction and learn something in return. Dialogic pedagogy often emphasizes the role of difference (or “otherness”) as a driver for learning (Asterhan et al., 2020).

Accordingly, *dialogic gaps* can be seen as emergent differences between incommensurable voices. A productive dialogue may require seeing reality from the eyes of the other, which can include overcoming incommensurability. Therefore, in contrast to Vygotskian approaches, the Bakhtinian dialogic gap assumes a bi-directional (or multi-directional) potential for change (Wegerif, 2008, 2011).

Research on collaborative learning consistently shows the importance of mutuality and mutual dependence in the success of collaborative learning (e.g., Cohen, 1994; Pearlstein, 2021; Schwartz, 1999; Webb et al., 2014). Some researchers foster mutuality by eliciting and/or grouping students to make sure voices differ at the onset of the interaction (Asterhan & Schwarz, 2007; Glachan & Light, 1982; Gutierrez-Santos et al., 2016; Schwarz et al., 2000; Schwarz & Asterhan, 2011). In mathematics, Schwarz et al. (2000) extended Glachan and Light's (1982) "two wrongs make a right" phenomenon by probing 10th graders' performance on a decimal fractions task and grouping them in dyads who exemplified different misconceptions. This Vygotskian "bug-correction" approach was successful, mainly when dyads applied behavior associated with productive group learning.

A main hypothesis of this article is that mutuality is the interpersonal set-up that carries the highest potential for learning from collaboration. This is because encompassing interpersonal set-ups assumes a Vygotskian zone of proximal development that is unidirectional, with increased chances that one (dominant) voice may not change in the interaction (Wegerif, 2008). In addition, similarity interpersonal set-ups at the onset of the interaction lack the generative power of a dialogic gap. Mutuality interpersonal set-ups require a dialogic gap—bi-directional at the onset of the interaction—so they may have the highest chances to yield learning in terms of developing students' thinking. Bakhtin-inspired dialogic pedagogy often focuses on individual differences that bear the potential for individual learning through interaction (e.g., Asterhan et al., 2020).

Dialogic pedagogy approaches advocate the decentralization of mathematics learning and instruction (Koschmann, 1996). According to this approach, fostering preferable students' interactions with other students and technology can result in learning mathematical ideas—not only bug correction or learning pre-determined curricular content, but also idiosyncratic conceptual learning (Barwell, 2016; Wegerif, 2008). Learning analytics systems can support this process, probing students' PESs to support dialogue.

For example, Gutierrez-Santos et al. (2016) developed a content-specific GFRM and used it to group students based on their solution to a dynamic visualization metaphor for the concept of a linear function. Since their focus was on developing the GFRM (rather than pedagogy), group formation in their study was based on optimizing the highest degree of difference between the students. This is close to mutuality, but not conceptualized as such, nor does it account for personal and interpersonal set-ups and how they influence learning. Gutierrez-Santos and colleagues' article inspired us to look deeper into group formation pedagogical considerations that can leverage learning theories.

We make a conceptual leap regarding a student's personal example space as an expression of his voice. Similarly, a voice in Bakhtin's dialogism is situated in a specific context—it is time and space dependent, experienced by an agent and regards

“something.” Both assume that learning is a change that happens in and from interaction with cultural agents—peers, teachers, and technology. We consider a PES recorded by a learning analytics system such as STEP to be an echo of a student’s voice, situated in the context of a specific task. We devised a study using example-eliciting tasks in mathematics to inform the design of a content-specific GFRM. Dialogic gaps may help develop students’ spaces regarding (mathematical) concepts such as the parabola. Thus, we hypothesize that mutuality interpersonal set-ups will be more effective for this learning type than others. We ask how different interpersonal set-ups influence developing students’ personal example spaces related to the parabola. And, which interpersonal set-up is the most effective one in doing so?

Methodology

With regard to participants, we were interested in developing PESs among students who had formally learned the topic of *quadratic functions*, as claimed by their mathematics teachers. Three such teachers, all from different Israeli high schools (in Hadera, Jerusalem, and Tel-Aviv), carried out five stages of the experimental protocol (see below). We included fifty students (out of ninety-four from the three classes) in the final analysis. Thirty-seven of the students were in 9th grade and thirteen were in 10th grade. Out of the final sample, forty-three completed all four tasks. The remaining seven did not complete task 2 (collaborative learning), so their data was marked as control (i.e., the “no treatment” condition).

Experiment Protocol

The experiment consisted of five stages over two 90-min meetings: *preparation* and *pre-test* in the first lesson; *group formation* between lessons, and then *collaborative learning* and *post-test* in the second lesson. All five stages (except for group formation) were conducted in computer laboratories. The STEP was open on every computer. Students were given usernames and passwords before the lesson. They were encouraged to use pen and paper for calculation.

The *preparation* stage (first meeting) began with students logging in and solving task 0 (see the next sub-section). The goal was to introduce students to STEP—logging into the system, navigating it, and submitting answers in the context of an example-eliciting task. After the students submitted three examples for task 0, their teachers used the students’ experience to discuss the idea of reducing the degrees of freedom of the problem (Pólya, 1945/1957) by choosing a second point in the graphical representation or a parameter in the algebraic representation (see the next sub-section). The preparation stage ended with a recap about the three forms of the quadratic function equation—polynomial, intercepts, and vertex forms—and how the parameters that comprise each form are reflected in the parabola graph (see the “Task analytics” section below).

The *pre-test* stage followed the preparation stage in the same lesson. In the pre-test, students solved task 1 as individuals. The *group formation* stage, administered

between the two meetings, started with a short recapitulation by the experimenter (first author) about the three group formation strategies—and similarity encompassment and mutuality. Teachers then used data on students' PESs to group them into pairs. Since this article is situated in the context of the design of a content-specific GFRM, we did not consider these groupings in our analysis because they were not accurate enough to be considered reliable. (In Abdu et al. (2019), we studied teachers' considerations for group formation in the context of example-eliciting tasks.)

This methodology, however, ensured variability of interpersonal set-ups (see later for further details about interpersonal set-up coding). We conducted a second 90-min meeting a week after the first one, which began with the fourth stage—*collaborative learning*. In this stage, every pair of students sat next to one computer. First, the students were asked to present their peers with their answers to task 1 in STEP. Next, every pair solved task 2 together. We gave no further instructions to the students. The last stage of the experiment was the *post-test*. In it, all students were asked to solve task 3, individually.

Tasks

The content we chose for the students to learn was about the parabola concept. Like many other mathematical concepts, the parabola has many symbolic and figural manifestations. For example, a parabola can intersect the x -axis in zero, one, or two places (for an elaborated list, see the next sub-section). As such, the parabola serves as an adequate context for probing for and developing students' PESs.

We developed four tasks for this study: task 0 as preparation, task 1 to probe for students' initial PESs (i.e., pre-test), task 2 for collaborative learning, and task 3 to probe students' PESs again (i.e., post-test). We did this by asking them to choose two points (one point in task 0) on an interactive diagram and create algebraic expressions of functions that pass through these points. Doing so required making a few more choices to reduce the degrees of freedom of the problem (Pólya, 1957), such as choosing an additional point in the graphical representation or selecting the value of a parameter in a particular algebraic representation.

In task 0, students were presented with a *GeoGebra*-based applet (Fig. 2, top left) and given the following instructions: “Below is an interactive diagram with one point that is colored in red. Choose a point with the ‘new point’ button and build a linear function that passes through that point. Repeat this three times to submit three different examples. Explain how these three examples are different.” In task 1, students were presented with a slightly different image (Fig. 2, top right) and given the following instructions: “Below is an interactive diagram with two points colored in red. Choose a pair of points with the ‘new points’ button and build with them a parabola. Repeat this three times to submit three different examples. Explain how these three examples are different?” The applet was pre-coded, so one of the given points was always somewhere on the x -axis, while the other was on the y -axis.

Task 2 was similar to task 1, but the applet was pre-coded, so both given points had the same y -value, while their x -values were different (Fig. 2, bottom left). Task 3 was similar to tasks 1 and 2, but included more choosing: one point was randomized,

Fig. 2 The *GeoGebra*-based applets for the four example-eliciting tasks. In all cases (except for the green point in task 3), the x - and y -values of the points' locations were integer numbers randomly chosen by STEP, in the range of $[-10, 10]$, upon clicking on the “new points” button

similarly to the case of task 0, and a student can select the other point by moving the slider to adjust the point's x - and y -values. Also, the description for task 3 included specific directions regarding the use of this diagram—instead of the instruction, “Choose a pair of points with the ‘new points’ button ...,” this part read, “Choose a point with the ‘new point’ button and adjust the slider to choose an additional point.” In other words, in task 3, we took off all of the restrictions so students could build any parabola they could choose.

Task Analytics

A student's thinking cannot be measured from her or his perspective, but can be synthetically measured. We were looking at aspects related to a parabola that have several characteristics. Variation in these characteristics marks the variety of possible answers for the tasks—possible personal example spaces. Accordingly, we programmed STEP to analyze students' examples, in relation to tasks 1, 2, and 3, according to four mathematical aspects: roots, extremum, algebraic form, and correctness (see Fig. 3 below). We refer to the manifestation of these aspects in students' examples as “characteristics.”

- **Roots**

A parabola may intersect the x -axis in zero, one, or two locations. These points are called “roots.” STEP can identify the number of roots of a submitted parabola. Task 1 presupposes that the parabola has at least one root (e.g., the

point $(-2, 0)$ in Fig. 2, task 1). Accordingly, in our experimental design, the case of zero roots was available only in task 3. (Task 2 was not included in the pre-post analysis.) As a result, we omitted from the data set seventeen examples with no roots submitted in the post-test.

- *Extremum*

A parabola can open upwards (i.e., has a *minimum*) or open downwards (i.e., has a *maximum*).

- *Form*

STEP can identify the algebraic form in which a parabola was inscribed. There are three such forms: (a) *polynomial*, $y = ax^2 + bx + c$; (b) *intercepts*, $y = a(x - x_1)(x - x_2)$; (c) *vertex*, $y = a(x - p)^2 + k$. The equations underlying all three forms give direct access to the parabola's extremum via the parameter a —a positive value for a means the function has a minimum and a negative value means it has a maximum. Except for a , each equation underlying each form gives direct access to a different parabola characteristic. The *polynomial* form, also known as the standard form, is the most commonly used one in Israeli schools. The parameter c in the polynomial form signifies the function's value when $x = 0$, signifying the parabola's intersection point with the y -axis. The x_1 and x_2 in the *intercepts* form signify the parabola's intersection point(s) with the x -axis—namely, its roots. The parameters p and k in the *vertex* form signify the parabola's extremum's respective x - and y -locations, i.e., the point (p, k) .

- *Correctness*

A submitted parabola that passes through two chosen points is marked as “correct” by STEP, which also marks cases in which a submitted parabola passes through one of the points.

Fig. 3 Aspects of the parabola and characteristics that relate to them

Aspects	Characteristics
Roots	0 1 2
Extremum	Minimum Maximum
Form	Polynomial Intercepts Vertex
Correct	1 2 3

Coding

We extracted data about students' performance in tasks 1, 2, and 3 from STEP into Excel sheets. We then coded this data to answer our research question: What interpersonal set-ups effectively develop students' personal example spaces? We coined the independent variable "interpersonal set-ups" to signify the relationship of the PES of one student to a peer within a single mathematical aspect. We termed the dependent variable "learning" to signify the characteristics that appeared in the post-test, but not in the pre-test, within a single mathematical aspect.

Coding for Interpersonal Set-ups: Independent Variable

Following the pre-test, teachers grouped students to learn together. Every PES exhibited by each student consisted of four mathematical aspects. We compared the mathematical characteristics of the examples submitted by the 43 participants (excluding the "no treatment" condition) with those submitted by their peers. These comparisons were coded into four *interpersonal set-ups*: encompassing, encompassed, mutuality, and similarity.

Encompassing and *encompassed* are the two sides of a relationship. We coded an interpersonal set-up as *encompassment* when the mathematical characteristics exhibited by one student (*giving*) include and exceed the characteristics exhibited by a second student (*receiving*). For example, the interpersonal set-up was coded as *encompassing* along the "roots" aspect when one student included parabolas with one root and two roots in her submission. In contrast, her peer only included parabolas with two roots.

Mutuality signifies interpersonal set-ups in which, within one aspect, the PESs of both students are different, and neither contains the other. Interpersonal set-ups were coded as mutual when each of the two students submitted at least one characteristic absent from her or his peer's examples—for example, a student who only submitted parabolas written in the polynomial form and a student who only submitted parabolas written in the vertex form.

Similarity refers to an interpersonal set-up in which two PESs are the same within one aspect. Interpersonal set-ups were coded as similar when neither student included an individual characteristic in her submission compared with her peer—for example, two students who both only submitted parabolas with maximal extrema.

Solving example-eliciting tasks with dynamic mathematics environments are typically open-ended and invite the learner for inquiry. Letting students decide what problems they would like to solve yields rich data about their PESs, but coding such rich data should be dealt with carefully. We found the need for such careful treatment in the correctness aspect of students' PESs. In its simplest form, correctness assumes an inherent encompassment between the manifestations of characteristics: A function that does not pass any point or only one point is an "incorrect example" and inferior to a function that passes through the two given points (correct). Correctness is thus different from the other three mathematical aspects.

There is no hierarchy between the three algebraic forms, number of vertices, or the two extrema (curriculum considerations or teachers' decisions may elicit a

Table 1 Independent variable: number of interpersonal set-ups ($n = 200$), with respect to mathematical aspects and interpersonal set-up types

	Extremum	Roots	Form	Correctness	Sum
Encompassing	7	11	6	7	31
Encompassed	10	12	5	4	31
Mutuality	0	1	9	22	32
Similarity	26	19	23	10	78
No treatment	7	7	7	7	28
Sum	50	50	50	50	200

Table 2 Number of added characteristics with regard to mathematical aspect and interpersonal set-up type. As could be seen, mutuality was especially effective in the case of correctness

	Extremum	Roots	Form	Correctness	Sum
Encompassing	2	4	4	6	16
Encompassed	2	7	2	4	15
Mutuality	0	1	5	24	30
Similarity	6	6	13	7	32
No treatment	0	3	2	4	9
Sum	10	21	26	45	102

hierarchy of competence by giving more attention to one topic or the other, but, from a mathematical standpoint, such a hierarchy does not exist). Also, correctness is influenced by explicit or implicit decisions—which algebraic form to use, how many roots the parabola will have in the x -axis, and what kind of an extremum. When students were asked to choose three examples, they practically chose three problems they preferred to solve (notably, almost 70% of the examples submitted in this experiment were correct). Therefore, the correctness aspect is an indicator of success in solving three *specific* problems chosen by the student. Our study has twelve such options—three algebraic forms, two possible root settings (one or two) and two possible extrema. Consequently, we refer to *correctness* as success in solving one of the twelve possible problems. Every correct example was treated as a separate characteristic.

Table 1 presents interpersonal set-ups' distribution with respect to mathematical aspects and interpersonal set-up types of the two-hundred interpersonal set-ups (fifty participants with four mathematical aspects per participant). Notably, the participants' number is lower than we wished (data collection was abrupt due to the COVID-19 pandemic). We regard this caveat in the limitation in the final section.

Coding for Learning in PES: Dependent Variable

We coded every student's PES for each of the four mathematical aspects and considered "learning" having taken place when a mathematical characteristic that did not appear in the pre-test showed up in the post-test. Learning in the correctness aspect was marked when a correct example did not appear in the pre-test, but did in the post-test. Table 2 summarizes the added mathematical characteristics, with respect

Table 3 Average addition of characteristics, with respect to mathematical aspect and interpersonal set-up type (n. c. stands for “no cases”)

	Extremum	Roots	Form	Correctness	Average
Encompassing	0.29	0.36	0.67	0.86	0.52
Encompassed	0.2	0.58	0.4	1	0.48
Mutuality	n. c.	1	0.56	1.09	0.94
Similarity	0.23	0.32	0.57	0.7	0.41
No treatment	n. c.	0.43	0.29	0.57	0.32

Table 4 A comparison between learning new mathematical characteristics between the five interpersonal set-ups

	Encompassing	Encompassed	Mutuality	Similarity
Encompassed	.84			
Mutuality	> .01*	> .01*		
Similarity	.43	.59	> .01*	
No treatment	.24	.33	> .01*	.53

*Stands for significant difference

to mathematical aspects and interpersonal set-up types. This analysis was not sensitive to whether the student submitted incorrect examples in task 1 or correct examples different from those submitted in Task 3. What we counted as learning was, thus, the additional competence as recorded with STEP.

Findings

This section shows how different interpersonal set-ups influence learning new mathematical characteristics from the pre-test to the post-test. We begin with a quantitative comparison of the influences of the different interpersonal set-ups and then follow up with a description of four learning trajectories recorded by STEP of four students who participated in this study. This section ends with an elaborate analysis of the learning sequence of one student.

Quantitative Analysis: a Comparison Between Interpersonal Set-ups

A total of 102 new mathematical characteristics were added between the pre-test and the post-test and distributed along 200 interpersonal set-ups. Thus, beyond interpersonal set-ups, the average addition was 0.51. Mutuality was the most effective interpersonal set-up (see Tables 3 and 4), specifically for the correctness and the number of roots (one case), but it did not seem superior in other cases. Being on the receiving end of an encompassment set-up was effective in the correctness characteristic, and not too far behind was being on the giving end of an encompassment set-up. Being on the giving end in an encompassment set-up was also the most beneficial in

the form characteristic. All interpersonal set-ups had a higher average of additions than the case of no collaborative learning (“no treatment”).

We conducted one-way ANOVA to test the differences among the five interpersonal set-ups. Samples were assumed to be taken from populations with equal variances based on Levene’s test ($W = .64$; $p = .63$). ANOVA indicated a significant difference between interpersonal set-ups ($F = 4.76$, $p < .01$). In a post hoc test, we used Fisher’s least significant difference to compare the five set-ups. The test results revealed significant differences ($p < .05$) between mutuality set-ups and all other set-ups. None of the other tests revealed a significant difference.

Examples of Learning Sequences

Interpersonal set-ups influence the chances that individual learning would occur. The results in the previous sub-section imply that mutuality set-ups are the most efficient for developing students’ PES. Next, we describe four particular cases of learning sequences by four students, followed by an elaboration for one of these cases (see the next sub-section). Every case (see Fig. 4) is used to illustrate the learning sequence by four students. These sequences were chosen because of their exemplary

Fig. 4 The learning of mathematical characteristics by four students throughout the experimental design. Each of the four examples represents a student’s progression along the four mathematical aspects of the tasks. Every example represents a typical interpersonal set-up. In task 1, we present the answers by the student and the peer of that student. Task 2 represents joint answers when solving together, while task 3 represents answers at the final stage of the experimental design

power regarding the four interpersonal set-ups: encompassing (A), encompassed (B), mutuality (C), and similarity (D).

Students A and B solved task 2 together. Their overall interpersonal relations were encompassing, where student A was supposed to “give” and student B was supposed to “receive.” In the post-test, both students showed minor additions—Student B employed a new form (vertices) that manifested in student A’s answer to task 1. Student A employed the intercepts form—which did not appear previously in his or his peer’s examples. In both cases, these attempts were incorrect. For student C (for whom we also provide a detailed analysis in the next sub-section), the overall interpersonal set-up related to her peer was mutual (see Table 5). Overall, four new characteristics were added to student A’s PES, including two new correct types. Student D created the same examples as his peer in the pre-test (i.e., similarly). The dyad created the same type of examples in task 2. Finally, student B did not manifest any new characteristics in task 3 (Table 5, bottom).

Student C’s Learning Sequence: a Case Study

In this sub-section, we describe the learning sequence of student C (pseudonym, Celia), bringing in snapshots starting from the examples she submitted in the pre-test, the results of her interaction with her peer (pseudonym, Paula) in task 2 and ending with the examples she submitted in the post-test.

Celia’s Examples in the Pre-test

Celia chose three similar examples in the pre-test—all of the parabolas were inscribed in vertex form, where the vertex chosen to be a minimum (extremum aspect) is the point on the x -axis (Fig. 5). Choosing this combination means that the solver knows what would be p and k in the vertex form, $y = a(x - p)^2 + k$, because p and k are the respective x - and y -locations of the parabola’s extremum (p, k). For example, in Fig. 5 middle, Celia chose the point (8, 0) as the minimum of the parabola and, accordingly, the equation became $y = a(x - 8)^2 + 0$, or $y = a(x - 8)^2$. The next step was to replace x and y with the parameters of the second point, in our example point (0, 4), and extract the value of the parameter a . Celia successfully solved this task in two examples (Fig. 5 left and right)—although the examples *looked* correct in the other case (Fig. 5 middle), both parabolas passed near the points on the y -axis, but did not cross them. Thus, only one example was considered correct by the STEP system in Celia’s answer to task 1.

The Examples in the Pre-test by Paula and Celia’s peer

Paula was Celia’s peer in task 2. Paula’s pre-test parabolas were inscribed in polynomial form, with two roots and a minimum (Fig. 6). Choosing this combination means that the solver knows what would be c in the polynomial form, $y = ax^2 + bx + c$, because c signifies the y -value of the point in which the parabola intersects the y -axis. For example, in Fig. 6 middle, the equation becomes $y =$

Table 5 Interpersonal set-ups and individual learning per the four cases in Fig. 4

Student	Interpersonal set-up by aspect				Overall interpersonal set-up relating to peer	Learning in regard to aspect				Total additions
	Roots	Extremum	Form	Correctness		Roots	Extremum	Form	Correctness	
A	>	=	>	>	>	0	0	1	0	1
B	<	=	<	<	<	1	0	1	0	2
C	><	=	><	>	><	1	0	1	2	4
D	=	=	=	=	=	0	0	0	0	0

Legend: *encompassing* ($>$); *encompassed* ($<$); *mutuality* ($><$); *similarity* ($=$)

Fig. 5 The three examples submitted by Celia in the pre-test

Fig. 6 The three examples submitted by Celia’s peer, Paula, in the pre-test

$ax^2 + bx - 4$. From this point, there are two possible routes for a solution. One would be to choose another point and, together with the point on the x -axis, create a system of two linear equations to solve for the parameters a and b . A second one would be to choose a value for the parameter a or b and then solve the equation for b or a . While all three examples submitted by Paula looked correct, while all three parabolas pass near the points on the y -axis, they actually do not cross them.

Comparing Celia and Paula’s Examples in the Pre-test

The difference between Celia’s and Paula’s examples can be analyzed considering the four mathematical aspects. Celia solved the same problem three times, succeeding only once. Paula solved another problem three times, succeeding in none. The correctness aspect was coded as encompassing, with Celia as the student with superior performance. Note that success was not measured here based on solving the same task. The extremum aspect was coded “similar”; roots and form aspects were both coded as “mutual” (Table 5, row C).

Celia and Paula's Joint Examples for Task 2

We identified two types of examples submitted by Celia and Paula together in Task 2. The first and second examples submitted (Fig. 7, left and middle) were similar—Celia and Paula chose a vertex form, minimum, and two roots based on two points on the two sides of the y -axis. These were also Paula's choices in the examples she submitted in the pre-test, but the two points in task 2 had the same y -value, which led to a different problem to be solved.

Parabolas are symmetrical and therefore the horizontal location of the parabola's extremum is at the mid-way between the two points colored in red. For example, in Fig. 7 (middle), the x -value of the parabola's extremum ($p = -3$) is mid-way between -10 and 4 (the distance between -10 and -3 equals the distance between 4 and -3). Once the value of p is known in vertex form, $y = a(x - p)^2 + k$, the solver can reduce degrees of freedom by choosing a value for the parameter a or k . Celia and Paula chose $a = 1$ in these two cases and then placed the x - and y -values of one of the points to extract the value of k . They succeeded in both cases.

The third example submitted by Celia and Paula included a hybrid approach. First, the two points they chose were placed symmetrically relative to the y -axis, a heuristic aimed to tap on the symmetry attribute of the parabola. This decision had ramifications for inscribing the function according to both vertex and polynomial forms. In the vertex form, this decision forced the x -value of the parabola's extremum p to be 0 , leading to the equation $y = a(x - 0)^2 + k$, which becomes a hybrid between the vertex and polynomial forms $y = ax^2 + k$. Here again, a solver can reduce the degrees of freedom by choosing a value for the parameter a or k . Celia and Paula chose the value $a = -1$, but failed to extract the value of k (note that if a were equal to 1 , the function would have passed through the two points).

Celia's Examples in the Post-test

In the post-test, all the parabolas submitted by Celia were correct, had a minimum, and had two roots. We identified two types of examples submitted by her. The first one (Fig. 8, left) was based on two points with the same y -value. One of these points,

Fig. 7 The three examples submitted together by Celia and Paula in task 2

Fig. 8 The three examples submitted together by Celia in the post-test

(0, 5), intersects with the y -axis. Celia chose to use a vertex form and, as in task 2, she identified the x -value location of the extremum ($p = 3$) and then chose a value for the parameter a ($a = 1$). Next, she extracted the value of k ($k = -4$) to come up with a final function. The second and third examples submitted by her were similar (Fig. 8, middle and right). In each example, one point intersects the y -axis, just like the first example.

Unlike the first example, Celia chose that this point would also be the minimum. As the case of the third example in task 2 (Fig. 7, right), using a vertex form in this case yields the simplified formula ($y = ax^2 + k$). After this alternation, this algebraic expression is in a polynomial form, where $b = 0$ and the parameter k replaces the parameter c . Now, Celia could choose any point on the Cartesian plane and place its x - and y -values into the formula to extract a value for a . In each of the two examples, the second point Celia chose is located in the first quadrant.

Four Mathematical Aspects of Celia's Learning Sequence

- *Roots*

In the pre-test, Celia only submitted examples with one root, while Paula only submitted examples with two roots (coded as mutuality). In task 2, the dyad submitted two examples with two roots and one example with no roots. In the post-test, Celia submitted three examples with two roots, thus incorporating a new characteristic to her examples.

- *Extremum*

All the examples related to Celia's learning sequence had a minimum (coded as similarity), except for one example in task 2. Thus, she did not incorporate any new characteristics into her examples.

- *Algebraic form*

In the pre-test, Celia only submitted examples in vertex form, while Paula only submitted examples in polynomial form (coded as mutuality). In task 2, the dyad submitted two examples in vertex form and one example in a hybrid between vertex and polynomial forms (we discuss below the implications of this finding to the design of our GFRM). In the post-test, Celia submitted one vertex form and two examples in a hybrid between vertex and polynomial forms. Thus,

she incorporated a new characteristic to her examples—STEP is programmed to identify a polynomial form when the algebraic expression has no parentheses, which was true for all tasks.

- *Correctness*

In the pre-test, Celia tried to solve one type of problem three times and was correct once and twice was close to giving accurate examples. Paula tried to solve another type of problem three times and submitted three incorrect examples—the parabola passed through one point in all three cases. This interpersonal set-up was coded as “encompassing.” In task 2, Celia and Paula submitted examples that combined the works of two of them. They submitted a third type of example twice and were correct, and a fourth type of example in which they were incorrect. Celia’s examples in the post-test included other characteristics that originated (a) in her example to the pre-test—the vertex form, (b) in Paula’s work, and later in task 2—a point on the y -axis that is also a minimum and two roots—and (c) in task 2 alone—a hybrid between polynomial and vertex forms and two points with the same y -value (Fig. 7, left). In sum, Celia submitted two new types of examples and was correct in both cases.

Discussion and Conclusions

Contemporary learning analytic-based modules can effectively optimize group formation (e.g., Maqtary et al., 2019; Meyer, 2009; Moreno et al., 2012). The invention of these modules can help teachers perform content-specific group formation, but only a handful of GFRMs are rooted in a content-specific pedagogical approach. This article explored pedagogical considerations for designing content-specific group formation recommendation modules (GFRMs) in terms of the combination among task, personal considerations, and the group formation approach (interpersonal considerations). We combined two learning-instruction principles—example-eliciting tasks (Yerushalmy et al., 2017; Yerushalmy & Olsher, 2020) and dialogic learning (Wegerif, 2011; Wegerif & Major, 2019)—and applied them in the context of a learning analytics system that can be used to support automated formative assessment in mathematics.

Developing a GFRM requires understanding the personal considerations for grouping—the data to collect about students (Maqtary et al., 2019). In this study, our content-specific objective was to foster students’ development of their PESs about the concept of parabola. We used example-eliciting tasks to collect data about their spaces. Such tasks are designed to probe for students’ PESs of specific (mathematical) content—or, at least, those answers that students choose to submit (Abdu et al., 2019; Olsher et al., 2016; Yerushalmy et al., 2017).

In the current study, students’ answers to three such tasks were analyzed along a pre-defined set of mathematical aspects to be considered when learning about parabolas—extremum, roots, form, and correctness. Breaking down a task to its fundamental mathematical aspects supported a fine-grained analysis of students’ learning sequences. Other tasks developed in our laboratory are geared to probe for students’ PESs regarding various mathematical concepts for many different curricular-focused

activities, such as asymptotes, linear functions, and triangles (Olsher et al., 2016; Yerushalmy & Olsher, 2020). For example, in a current task, we ask students to create three different examples of triangles that will be analyzed via several interweaving aspects, regarding angles (right, acute, or obtuse), line segments (equilateral, isosceles, or scalene), and segment-positioning (vertical, horizontal, or one vertical and one horizontal).

As manifested in the diverse interpersonal set-ups (Table 1) and learning (Table 2), the examples submitted in tasks 1, 2, and 3 were diverse. We exemplified this diversity in the work of four students who were each engaged in a different interpersonal set-up (Table 5). This diversity of these PESs indicates the potential of example-eliciting tasks and accompanied analytics in fostering content-specific idiosyncratic learning of specific content (Olsher et al., 2016; Yerushalmy & Olsher, 2020). Some students' solutions to the example-eliciting tasks in this experiment unveiled an ambiguity in the analytics—cases in which the form of the function was a hybrid between vertex and polynomial forms (such as in Celia and Paula's joint examples for task 2; see also Table 5, row C and D). Designing educational technology is a cyclic process that includes iterations of implementations and correction (Bakker, 2018), and the design of content-specific GFRMs is no different.

Developing a content-specific GFRM requires studying the influence of interpersonal relations on the effectiveness of individual learning. We operationalized such relations in this experiment using distinctions between what we called “interpersonal set-ups”—encompassment, mutuality, and similarity—along with the four mathematical aspects of the parabola concept. Our findings indicate that the most effective interpersonal set-up for learning about the parabola in the specific settings of this experiment was “mutuality” (Table 4)—when every student who joined the group had manifested a singular characteristic when working alone (e.g., Gutierrez-Santos et al., 2016). These findings support the hypothesis that mutuality set-ups are superior when the content-specific objective fosters students' development of their PES (specifically, in our case, the concept of parabola). However, it is worth stressing that this probing methodology for PESs through example-eliciting tasks does not consider students' level of responding to the task (Kontorovich et al., 2012) and cannot be regarded as an accurate reflection of what students know or learned in the process.

A personal example space in this article's context reflects what a student chooses to submit as part of her or his repertoire of available examples and ways to construct them. Notably, most of the interpersonal set-ups coded mutual-regard correctness and most of the correctness interpersonal set-ups were effective. The effectiveness of mutuality in respect to form was good, but not better than other interpersonal set-ups. The only other case was mutuality interpersonal set-up regarding roots and interaction that was fruitful for learning (see Table 3). Thus, learning to create new correct examples was best in mutuality interpersonal set-ups and was indeed the main focus of improvement in every interpersonal set-up, but not at the same size and scale of the phenomena for mutuality ones.

The theory of dialogic thinking and learning (Bakhtin, 1984; Wegerif, 2011) helped us to contemplate about content-specific, mutuality group formation as a process in which teachers used technology to elicit and probe for students' voices

(personal example spaces) and to group them based on a potential dialogic gap between them (mutuality set-ups). The designer of a content-specific GFRM defines the characteristics relevant to content learning to be probed. As a socio-cultural, ontological entity, a dialogue is always situated and contextualized.

Dialogic gaps—the incommensurable differences between personal example spaces—vary in their magnitude (e.g., how many characteristics apart?) and type (i.e., which characteristics are compared?). In the current study, the similarity was regarded as the lack of a dialogic gap. The interpersonal set-up for extrema between Celia and Paula was coded as similar, but the y -locations of their solutions to task 1 were different and incommensurable. An example-eliciting task is a socio-cultural entity that can be calibrated to calculate pre-determined characteristics. Indeed, this study shows that a dialogic gap at the onset of the interaction related to the four characteristics we chose to address as designers had the highest chance to yield individual learning out of collaborative learning situations.

A dialogic outlook on content-specific grouping predicts that lack of dialogic gap would restrain the pedagogical potential for fostering students' development of their PESs about specific content. Accordingly, similarity interpersonal set-ups were framed in this study at the level of particular mathematical aspects. The results illustrated in Table 4 further indicate that similarity was the least effective interpersonal set-up (non-significant). The extremum in Celia's learning sequence is a case in point—all of the examples in her and her peer's pre-test had a minimum, one incorrect example had a maximum in their joint solution, and all of her examples had a minimum in the post-test.

The absence of a dialogic gap in this case co-occurred with the absence of individual learning. Student D's learning sequence (see Table 5) is another illustration of the (in-) effectiveness of similarity grouping in the case of an example-eliciting task—grouped with a peer who submitted similar examples in the pre-test, along with all four mathematical aspects—and did not manifest any change of his PES in later tasks. Designers of content-specific GFRMs can leverage this insight in focusing on mutuality considerations for group formation when the pedagogical objective is to support students' learning of their PESs. Different combinations of tasks, personal considerations, and interpersonal considerations may yield different individual learning outcomes.

Teachers' reasons for grouping can vary according to the learning goals they set up. Many tasks focus on mastery, in which learning objectives are “to know something” and, in these cases, the most efficient interpersonal set-up will be that of encompassment. However, similarity may be the most fitting group formation when the instructional goal is to provide students with adaptive support (e.g., ability grouping; see Park & Datnow, 2017). Which interpersonal set-ups are efficient for which type of content-specific learning? And with regard to which aspects?

This theory-driven study design taps on previous studies on collaborative learning, which stress that taking the time to think outside of the confirmative environment of a dialogue (co-operation) may create an opportunity for a dialogic gap to open. When two voices acknowledge the gap, agents might strive to overcome it—by listening to each other and maybe by accepting the other's perspective. These attempts of a student to take others' perspectives could result in

the development of her or his voice (or at least see the other's perspective) (Abdu et al., 2021). In the current study design, we operated such interplays between individual and collaborative learning. In a classroom advocating decentralized learning and instruction (Koschmann, 1996), these interplays can be broadened to longer learning sequences in which students learn alone with STEP and also when grouped with other students together based on the PESs they submit.

Fostering a dialogic gap at the onset of the interaction may be effective in learning sequences like the one in this study. Nevertheless, the effectiveness of a collaborative learning situation also depends upon other factors, such as the support given to learners (e.g., Dekker & Elshout-Mohr, 2004) and how students interact when they learn together. Collaborative learning is effective when the learners are all engaged, ask questions, listen to each other, and think either together or alone (e.g., Barron, 2003; Schwarz et al., 2000; Webb et al., 2014). These are all confounding factors that may or may not have appeared when students learned in groups in the current study and may sway this study's results one way or another. Thus, further work should focus on a microanalytic analysis of multimodal learning interactions for de facto interpersonal set-ups to characterize how different interpersonal dynamics and teacher interventions lead to the development of PESs. The other side of that coin regards macro-processes of longitudinal learning gains of dialogic versus monologic GFRM considerations.

Further studies can address another limitation of the current study—the relatively small number of participants and limited attention to comparing interpersonal set-ups at the PES level. Researchers who wish to conduct such a study will be faced with the challenge of developing a measure that captures both the size (how many singular new characteristics are manifested by individual students) and the direction (how these characteristics are distributed among these students) of the dialogic gap between voices. This measure of the gap can inform group formation and be used as an independent variable that influences the development of example spaces.

In the future development of content-specific GFRMs that are meant to develop students' personal example spaces, teachers can define the (mathematical) aspects they would like to develop, design tasks that can elicit such student spaces, and build on mutuality of perspectives (see also the work cited above by Schwarz and colleagues on argumentative and dialogic designs). One can do this kind of group formation for aspects of concepts for which engineers can extract data. Hypothetically, teachers could discern students' thoughts or beliefs about certain historical events, people, or political views, and group them in ways that may help them learn from each other. We hope this study will inspire content-specific experts to develop their versions of content-specific GFRMs.

Acknowledgements This research was supported by the Yuval Ne'eman fellowship from the Israeli Ministry of Science and the Israeli Scientific Foundation (ISF) (grant 147/18).

Declarations

Competing Interests The authors declare no competing interests.

References

- Abdu, R., Olsher, S., & Yerushalmy, M. (2019). Towards automated grouping: unraveling mathematics teachers' considerations. In B. Barzel et al., (Eds), *Proceedings of the 14th International Conference on Technology in Mathematics Teaching* (pp. 147–154). Essen, Germany: ICTMT.
- Abdu, R., van Helden, G., Alberto, R., & Bakker, A. (2021). Multimodal dialogue in small-group mathematics learning. *Learning, Culture and Social Interaction*, 29, (#100491).
- Amara, S., Macedo, J., Bendella, F., & Santos, A. (2016). Group formation in mobile computer-supported collaborative learning contexts: A systematic literature review. *Journal of Educational Technology & Society*, 19(2), 258–273.
- Asterhan, C., & Schwarz, B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, 99(3), 626–639.
- Asterhan, C., Howe, C., Lefstein, A., Matusov, E., & Reznitskaya, A. (2020). Controversies and consensus in research on dialogic teaching and learning. *Dialogic Pedagogy*, 8.
- Bakhtin, M. (1984). *Problems of Dostoevsky's Poetics* (Trans. C. Emerson). University of Minnesota Press.
- Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge.
- Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12(3), 307–359.
- Barwell, R. (2016). Formal and informal mathematical discourses: Bakhtin and Vygotsky, dialogue and dialectic. *Educational Studies in Mathematics*, 92(3), 331–345.
- Berland, M., Davis, D., & Smith, C. (2015). AMOEBA: Designing for collaboration in computer science classrooms through live learning analytics. *International Journal of Computer-Supported Collaborative Learning*, 10(4), 425–447.
- Borges, S., Mizoguchi, R., Bittencourt, I., & Isotani, S. (2018). Group formation in CSCL: A review of the state of the art. In A. Cristea, I. Bittencourt, & F. Lima (Eds.), *Higher education for all. From challenges to novel technology-enhanced solutions* (pp. 71–88). Springer.
- Buber, M. (1923). *I and Thou*. Charles Scribner's Sons.
- Cohen, E. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1–35.
- Connor, C., Morrison, F., Fishman, B., Crowe, E., 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.
- D'Angelo, C., Roschelle, J., Bratt, H., Shriberg, L., Richey, C., Tsiartas, A., & Alozie, N. (2015). *Using students' speech to characterize group collaboration quality*. Poster presented at the Computer-supported Collaborative Learning Conference.
- 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.
- Erkens, M., Bodemer, D., & Hoppe, U. (2016). Improving collaborative learning in the classroom: Text mining-based grouping and representing. *International Journal of Computer-Supported Collaborative Learning*, 11(4), 387–415.
- Glachan, M., & Light, P. (1982). Peer interaction and learning: Can two wrongs make a right? In G. Butterworth & P. Light (Eds.), *Social cognition: Studies of the development of understanding* (pp. 238–262). University of Chicago Press.
- Goldenberg, P., & Mason, J. (2008). Shedding light on and with example spaces. *Educational Studies in Mathematics*, 69.
- Gutierrez-Santos, S., Mavrikis, M., Geraniou, E., & Poulouvasilis, 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.
- 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.
- Holmberg, K. (2019). Formation of student groups with the help of optimisation. *Journal of the Operational Research Society*, 70(9), 1538–1553.

- Isotani, S., Inaba, A., Ikeda, M., & Mizoguchi, R. (2009). An ontology engineering approach to the realization of theory-driven group formation. *International Journal of Computer-Supported Collaborative Learning*, 4(4), 445–478.
- Johnson, D., & Johnson, R. (2002). Learning together and alone: Overview and meta-analysis. *Asia Pacific Journal of Education*, 22(1), 95–105.
- Konert, J., Burlak, D., & Steinmetz, R. (2014). The group formation problem: An algorithmic approach to learning group formation. In C. Rensung, S. de Freitas, T. Ley, & P. Muñoz-Merino (Eds.), *Open learning and teaching in educational communities: 9th European Conference on Technology Enhanced Learning* (pp. 221–234). Springer.
- 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.
- Koschmann, T. (1996). Paradigm shifts and instructional technology. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 1–23). Lawrence Erlbaum Associates.
- Lou, Y., Abrami, P., Spence, J., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66(4), 423–458.
- 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), 169–190.
- Martinez-Maldonado, R. (2019). A handheld classroom dashboard: Teachers' perspectives on the use of real-time collaborative learning analytics. *International Journal of Computer-Supported Collaborative Learning*, 14(3), 383–411.
- Meyer, D. (2009). OptAssign: A web-based tool for assigning students to groups. *Computers & Education*, 53(4), 1104–1119.
- Moreno, J., Ovalle, D., & Vicari, R. (2012). A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. *Computers & Education*, 58(1), 560–569.
- 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.
- Park, V., & Datnow, A. (2017). Ability grouping and differentiated instruction in an era of data-driven decision making. *American Journal of Education*, 123(2), 281–306.
- Pearlstein, J. (2021). Experiential exercise in team formation in the capstone: Providing students with the information to make good team choices. *Journal of Management Education*, 45(4), 627–651.
- Pólya, G. (1945/1957). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
- Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.
- Schwarz, B., & Asterhan, C. (2011). E-moderation of synchronous discussions in educational settings: A nascent practice. *The Journal of the Learning Sciences*, 20(3), 395–442.
- Schwarz, 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.
- Schwarz, B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right ... if they argue together! *Cognition and Instruction*, 18(4), 461–494.
- Schwarz, B., Prusak, N., Swidan, O., Livny, A., Gal, K., & Segal, A. (2018). Orchestrating the emergence of conceptual learning: A case study in a geometry class. *International Journal of Computer-Supported Collaborative Learning*, 13(2), 189–211.
- Schwartz, D. (1999). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 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.
- Slavin, R. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293–336.
- Teo, P. (2019). Teaching for the 21st century: A case for dialogic pedagogy. *Learning, Culture and Social Interaction*, 21, 170–178.
- Tsovaltzi, D., Weinberger, A., Schmitt, L., Bellhäuser, H., Müller, A., Konert, J., Röpke, R., Eshuis, E., Eysink, T., Erkens, M., Manske, S., Hoppe, U., Bodemer, D., Sankaranarayanan, S., Dashti, C., Bogart, C., Wang, X., Sakr, M., Hilton, M., & Rosé, C. (2019). Group formation in the digital age: Relevant characteristics, their diagnosis, and combination for productive collaboration. In K. Lund,

- G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *Proceedings of the 13th International Conference on Computer-Supported Collaborative Learning* (Vol. 2, pp. 719–726). International Society of the Learning Sciences.
- van Leeuwen, A. (2015). Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload. *Journal of Learning Analytics*, 2(2), 138–162.
- van Leeuwen, A., & Janssen, J. (2019). A systematic review of teacher guidance during collaborative learning in primary and secondary education. *Educational Research Review*, 27, 71–89.
- Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology*, 14(3), 293–305.
- Vygotsky, L. (1978). *Children's mental development in the process of learning*. In *Moscow–Leningrad. USSR: Uchpedgiz*.
- Wang, D., Lin, S., & Sun, C. (2007). DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, 23(4), 1997–2010.
- Webb, N. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79(1), 1–28.
- Webb, N., Franke, M., Ing, M., Wong, J., Fernandez, C., Shin, N., & Turrou, A. (2014). Engaging with others' mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning. *International Journal of Educational Research*, 63, 79–93.
- Wegerif, R. (2008). Dialogic or dialectic? The significance of ontological assumptions in research on educational dialogue. *British Educational Research Journal*, 34(3), 347–361.
- Wegerif, R. (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity*, 6(3), 179–190.
- Wegerif, R., & Major, L. (2019). Buber, educational technology, and the expansion of dialogic space. *AI & Society*, 34(1), 109–119.
- Wise, A., & Schwarz, B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4), 423–467.
- Yerushalmy, M., & Olsher, S. (2020). Online assessment of students' reasoning when solving example-eliciting tasks: Using conjunction and disjunction to increase the power of examples. *ZDM: Mathematics education*, 52(5), 1033–1049.
- Yerushalmy, M., Nagari-Haddif, G., & Olsher, S. (2017). Design of tasks for on-line assessment that supports understanding of students' conceptions. *ZDM: Mathematics education*, 49(5), 701–716.
- Zheng, Y., Li, C., Liu, S., & Lu, W. (2018). An improved genetic approach for composing optimal collaborative learning groups. *Knowledge-Based Systems*, 139(C), 214–225.
- Zheng, Z., & Pinkwart, N. (2014). A discrete particle swarm optimization approach to compose heterogeneous learning groups. In *Proceedings of the IEEE 14th International Conference on Advanced Learning Technologies* (pp. 49–51). IEEE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.