Multimodal Dialogue in Small-Group Mathematics Learning

Abdu, Rotem; van Helden, G.; Alberto, Rosa; Bakker, Arthur

DOI
10.1016/j.lcsi.2021.100491

Publication date
2021

Document Version
Final published version

Published in
Learning, Culture and Social Interaction

Learning, Culture and Social Interaction 29 (2021) 100491

Contents lists available at ScienceDirect

Learning, Culture and Social Interaction

journal homepage: www.elsevier.com/locate/lcsi

Multimodal dialogue in small-group mathematics learning

Rotem Abdu [a], Gitte van Helden [b], Rosa Alberto [b], Arthur Bakker [b,*]

[a] University of Haifa, Haifa, Israel
[b] Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Dialogue
Multimodality
Learning
Small group
Mathematics lesson

ABSTRACT

In this paper, we combine dialogic and embodied theories of learning to create a unified analytic lens. Embodied cognition is a theoretical approach operating under the premise that thinking and communication are multimodal activities. Under this premise, dialogue between learners needs to be conceptualized using a multimodal lens. We identify multimodal voices as speech and movement bundles situated within a learning context and describe a phenomenon that we call *Multimodal Dialogue* – multimodal interaction between different multimodal voices. To demonstrate this phenomenon, we analyze a learning sequence by two third-grade students who participated in a mathematics lesson aimed to foster embodied learning of proportion. Our analysis zooms in on the phenomenon of a multimodal voice as a speech-and-movement bundle situated within a learning context. We further show how multimodal dialogic gaps – differences between multimodal voices within and between modalities – drive communication and eventual changes in voices.

1. Introduction

This paper explores the value of combining dialogic and embodied (multimodal) approaches to analyze small-group mathematics learning. Our premise is that people learn in and through multimodal (e.g., speech, gesture, movement) interaction with the environment. Accordingly, both dialogic (Cresswell & Teucher, 2011; Wegerif, 2011) and embodied (Flood, 2018; Radford, 2009) perspectives consider learning a transient, culturally embedded, and situated activity.

Dialogic pedagogy theories conceptualize learning as a personal meaning-making process situated in interpersonal interactions (Asterhan et al., 2020; Buber, 1923; Wegerif, 2011) – usually focusing on speech modality without explicit consideration of the body's role in learning. Embodied learning theories focus on personal learning situated in multimodal interactions with cultural artefacts and instructors (e.g., Abrahamson, 2009; Flood et al., 2020; Lindgren & Johnson-Glenberg, 2013; Zhou et al., 2019), but rarely with explicit consideration of dialogue. Only a few studies have used dialogic and multimodal approaches in tandem to analyze teacher-led interactions (e.g., Bridges et al., 2020; Hennessy, 2011). Even less attention is paid to multimodal dialogue between learners in small groups. We propose that the intersection between embodied and dialogic perspectives for group learning can enhance theoretical understanding and can have implications for teaching and design. We use the term *multimodal dialogue*[1] to signify a phenomenon in which interactions occur within and across modalities. To explore this phenomenon's potential value for learning, we ask: *how can a*

* Corresponding author.
E-mail address: a.bakker4@uu.nl (A. Bakker).

[1] We use the term multimodal slightly differently than Taylor, Bouwhuis, and Neel (2000), who focus less on learning and more on the human-computer-interface aspects of dialogue.

https://doi.org/10.1016/j.lcsi.2021.100491
Received 6 March 2020; Received in revised form 21 December 2020; Accepted 31 December 2020
Available online 15 February 2021
2210-6561/

multimodal dialogue be manifested in group learning of mathematical content?

Section 2 presents and combines the two theories we chose to employ in our analysis, namely embodied design (Abrahamson, 2009) and the dialogic theory of thinking and learning (Wegerif, 2011). To concretize the theoretical potential of using these two theories in tandem, we describe a set-up (Section 3) and a learning sequence (Section 4) of a dyad who participated in a design study aimed to promote multimodal learning of proportion. The dyad learned about proportion through multimodal interaction with a designated embodied learning tool, and with each other. Section 5 sketches some implications for combining embodied and dialogical perspectives for design research.

2. Theoretical framework

Multimodal dialogue is a phenomenon that can be described using dialogic and embodied learning theories. This chapter elaborates on each of these two theoretical perspectives and follows with a description of a short proposal for a unified analytic lens to be developed and exemplified later in the findings section.

2.1. The dialogic theory of thinking and learning

Bakhtin's (1984) ideas of dialogue inspired educators to consider learning as change happening in and through interaction with other people and/or cultural agents (e.g., Marjanovic-Shane et al., 2019; Teo, 2019; Wegerif & Major, 2019). Bakhtin-inspired pedagogical approaches usually focus on opportunities that are given to learners to express themselves and to interact with others through verbal activity. Central to this approach is the role of differences between learners as drivers of learning – as explicated by E. Matusov, "we're hunting for the difference, for the otherness" (Asterhan et al., 2020, p. 10).

We focus on dialogic theories that help in understanding group learning – and specifically on three ontological entities from Wegerif (2006, 2011) dialogic theory of thinking and learning: voice, dialogic gap, and interaction. A *voice* is a perspective about "something," constrained by the environment and context in which it is expressed. Voices are thought or expressed by humans, but also by cultural artefacts (Wegerif & Major, 2019). A voice that is situated within social settings usually is in flux, since it may change whenever it interacts with other voices embedded in other personal or cultural discourses (Akkerman & Niessen, 2011; Bakhtin, 1984). Learning happens when voices change within or as a result of interaction with other voices. When two voices relate to something in two incommensurable ways, their interaction may foster change in one or two of them (Sfard, 2019). Interaction between similar voices can hardly foster change – and can be considered monologue (Bakhtin, 1984). Thus, dialogue requires a difference between the voices participating in the interaction – a *dialogic gap* (Wegerif, 2011). Dialogue happens when and where two or more distinct voices interact (Bakhtin, 1984). Engaging in dialogue means *interacting* with the other in a way that is open to the possibility that one's voice may change (Wegerif, 2011). A dialogue can promote learning when interlocutors acknowledge the dialogic gap between them (or their otherness, Buber, 1923; Wegerif & Major, 2019), explicate their thoughts to others (Sfard, 2019), listen to the other's voice (Arcavi & Isoda, 2007; Buber, 1923) and attempt to rearticulate and draw on other voices' logic (Sfard, 2019; Wegerif, 2011).

When persons engaged in a dialogue manage to overcome incommensurability and see the reality from the others interlocutors perspective, they may change – whether in a complete shift of perspective or being able to maintain two perspectives all at once (Sfard, 2019). If we agree that thinking is communicating with oneself (Sfard, 2008), then dialogic thinking may happen when an individual maintains two incommensurable voices and thus can see and experience reality from these two (or more) perspectives (Wegerif, 2011). For example, one person could maintain the perspective that the sun sets in the west when navigating, and the perspective that earth spins around itself when trying to explain to a child how the solar system works.

Learning in and through a Bakhtinian dialogue is a process driven by gaps and interactions. However, classrooms are places in which it may be harder to reach authentic Bakhtinian dialogues between peers – students are not always encouraged to learn in small groups, form idiosyncratic perspectives on a specific topic, or argue with peers (Alexander, 2008; Matusov et al., 2019). Moreover, dialogues between peers for learning specific (mathematical) content are prone to become monologues: Sometimes ideas converge, people agree or refrain from talking about differences of perspectives, and the dialogue dies (Marjanovic-Shane et al., 2019). Productive learning dialogues require students to mix individual thinking (to form voices) and interaction with others (to change voices). Students can be guided to take time off the interaction to think alone when the interaction converges to a monologue, and by that develop idiosyncratic perspective before and within the interaction with others (Abdu & Schwarz, 2020; Schindler & Bakker, 2020; Wit, 2006).

Dialogic theories mainly focus on interaction through speech modality with less attention to the material and embodied (Hetherington et al., 2018). In education, this focus on speech modality may be reasonable because speech usually carries rich semiotic information. Nevertheless, thinking and communication are increasingly acknowledged to be multimodal (Hetherington et al., 2018; Norris, 2004; Radford et al., 2017; Steier et al., 2015; Yoon et al., 2011). What would more explicit attention to modalities other than speech teach us about dialogue in group learning?

2.2. Embodied cognition

New connections between movement sciences and conceptual learning sparked new approaches to human cognition referred to as *embodied cognition* (Galetzka, 2017; Glenberg, 2010; Wilson & Golonka, 2013). From embodied cognition perspective, thinking and learning are multimodal (Anderson et al., 2012; Galetzka, 2017; Glenberg, 2010). Subsequent studies in mathematics education unravelled the centrality of the body in various learning processes (e.g., Abrahamson & Sánchez-García, 2016; Arzarello & Sabena, 2014; Hall & Nemirovsky, 2012; Radford et al., 2017). This understanding of the centrality of the body in thinking and learning ignited the development and evaluation of tools, tasks, and instructional contexts to foster and analyze embodied learning (e.g., Abrahamson & Bakker, 2016; Bridges et al., 2020; Lee, 2016; Sinclair et al., 2016; Zhou et al., 2019; Zohar & Levy, 2019).

In this paper, we focus on one research program instigated by Abrahamson and his colleagues called embodied design, which leverages embodied cognition theories for learning mathematics (Abrahamson, 2009; Abrahamson, 2015; Duijzer et al., 2017). Learning mathematical concepts in embodied designs typically happens *in* and *through* multimodal interaction with software modules called Mathematics Imagery Trainers (hereafter called "Trainers"), embedded in Wii, tablets, or touchscreens (Abrahamson, 2015). Interacting with *Trainers* elicits body movement – mostly hands – potentially grounding mathematical concepts in sensorimotor activity, such as proportion (Duijzer et al., 2017), the Cartesian field (Abrahamson & Bakker, 2016), or parabola (Shvarts & Abrahamson, 2019).

Specifically, in the Trainers for proportions, students are asked to move crosshairs or bars until the background of the interface or the manipulated bars become green (see Figs. 1, 4, 5, and 8 for the case of proportion).[2] These "green locations" are set up to when bars respective heights above the monitor base are related according to a particular proportion (e.g., 1:2 or 2:3). For example, for a 1:2 proportion, the bars will become green when they are 2 cm and 4 cm above the base. The bars will remain green if the user will move both index-fingers up or down while maintaining that proportion, say to 4 cm and 8 cm above the base. Note that, in so doing, the vertical interval (the difference) between the bars grow as the hands ascend and shrink when the hands descend. The Trainer has several screen overlays – disciplinary frames of references – such as a grid and numerals. The experimenter ideally introduces these after the user has demonstrated critical performance criteria. Students use these overlays initially to enact better or explain their solution strategy, such as validating that the interval indeed grows.

When students learn with a Trainer, they first need to develop coordination between their hands that grounds mathematical frames of references. Next, the students interact with the Trainer in search of an invariant property of the objects they interact with (Abrahamson & Abdu, 2020). For example, in the above example, a 1:2 proportion between the two bars is an invariant feature of the green locations. A paradigmatic interaction sequence towards discovering bimanual coordination congruent with the goal state would first involve mainly movement (Abrahamson & Trninic, 2015) as the student (a) explores the environment and first positions the hands incorrectly (red feedback); then (b) stumbles upon a correct position (green); and (c) raises both hands to maintain a fixed interval between them (red); and then (d) corrects the position with either hand (green). Through trial, error, and correction, learners do develop task-effective solution strategies (see Fig. 1). Elicitation of speech modality would be usually done by an experimenter in the room, asking the student to explain their strategy. There are multiple ways in which students move to solve this coordination task, and a variety of ways in which they attend to the problem space (for example, see also Figs. 4, 5, and 8). For example, students often refer to the interval and explain, "the higher my hands go, the bigger the interval" or "my right moves twice as fast as my left" (Abrahamson et al., 2014; Duijzer et al., 2017). As such, the students increasingly see, move, and talk in ways in line with normative practices.

By solving a movement task, students develop new coordination between modalities that unfolds as mathematical meanings. Students first learn to move in a new way and then – often as a result of researcher elicitation – analyze and articulate their movements by adopting mathematical words and gestures to enhance the enactment, explanation, or evaluation of their strategy (Duijzer et al., 2017). The movements and gestures users employ in the interaction with Trainers are sets of signs used to communicate a perspective in dialogue with a cultural voice. Arzarello, Sabena, and their colleagues (Arzarello & Sabena, 2014; Arzarello et al., 2009) use the term *semiotic bundle* to signify a transient system of signs (about something) explicated in one or more modalities. A semiotic bundle may or may not bare consistency between modalities. For Arzarello and Sabena (2014), multimodal thinking happens when a learner compares and tries to consolidate between sets of signs on different semiotic planes – such as gestures, mathematical signs, or words used. Similar to the idea of the dialogic gap as a driver of change, multimodal learning happens when semiotic meanings across modalities align.

2.3. Multimodal dialogue

Dialogic thinking theory can be elaborated to include learning as a multimodal phenomenon, and the embodied design perspective can lean on dialogic thinking theories to analyze and design for peer-to-peer interactive movement. A *multimodal dialogue* thus occurs when two distinct multimodal voices interact. In Fig. 2, we propose a model of a multimodal dialogue situation within/between two multimodal voices. For illustrative purposes, we restricted voice to two modalities — speech and movement — bundled together (in Figs. 2, 6-7, we denote multimodal bundles as two curved dotted arrows forming a circle). Note that the number of modalities can be extended to any relevant number. The model includes all the possible interactions within/between modalities (the vertical aspect of Fig. 2) and within/between voices (the horizontal aspect of Fig. 2). Note that this model does not account for the type of interactions between modalities. We develop this model in the remaining of this paper.

In the next section, we present the context for a case study we use to investigate *how a multimodal dialogue can be manifested in group learning of mathematical content*. By answering this research question, we hope to learn how multimodal dialogues may inform learning and instruction – of mathematical content in our case. Let us now describe the setting and context of a case study in which we will illustrate three multimodal dialogue ingredients: a multimodal voice, a multimodal dialogic gap, and multimodal interaction.

[2] All the figures in this paper were made by Gitte van Helden.

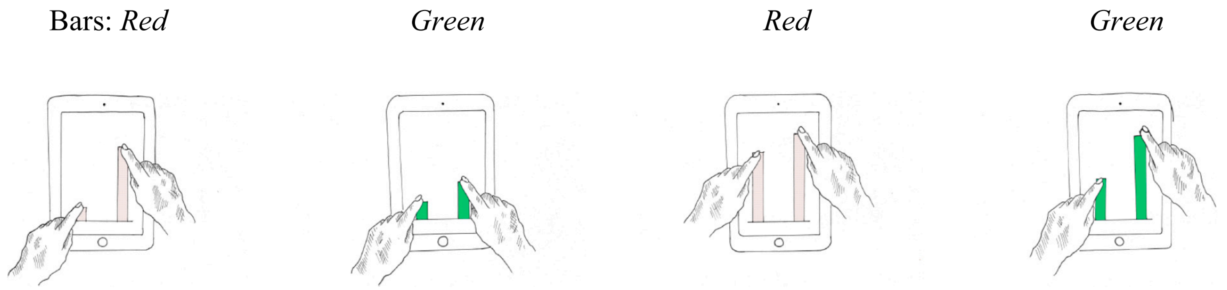

Fig. 1. The Mathematics Imagery Trainer for Proportion (in short: Trainer): a possible activity sequence (all drawings have been made by Gitte van Helden).

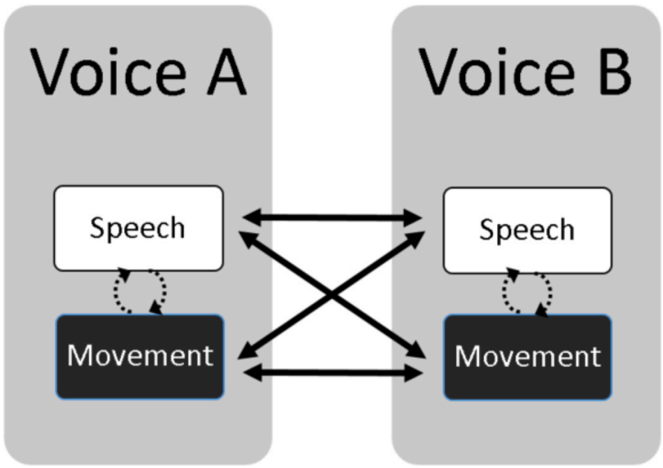

Fig. 2. A model of a multimodal dialogue situation. Consistent with our analysis, we focus on movement and speech modalities.

3. Setting and context

To explore the potential value of combining dialogic thinking and embodied design approaches, we chose a case study from a lesson on proportionality, in which the Trainer was combined with traditional on-paper learning tasks. This lesson was a part of design research (Bakker, 2018; Collins et al., 2004) aimed to instil embodied learning in mathematics classrooms. Designers and theoreticians of embodied design provide evidence for the effective multimodal tool-based learning of mathematical concepts. However, most of the empirical work on the topic of embodied design until now was still done in laboratory settings with a loose connection to the mathematics classroom.

Learning in embodied design studies framework was usually done in one-on-one settings, where expert researchers prompted the student to explain their movements (e.g., Shvarts & Abrahamson, 2019). In a classroom, the teachers undivided attention to the learning of every student, along multiple modalities, is not an option. This caveat can be overcome with small *group learning*, assuming that the interaction between peers may elicit learning through dialogue, within and across modalities: a multimodal dialogue. The lesson design mainly followed the task sequence used outside classrooms (e.g., Duijzer et al., 2017), but some adjustments were made to make the lesson suited for the classroom context. A practical decision was to let students work in pairs so that they could engage in a dialogue without the continuous presence of a tutor.

The students at the centre of this case are two third-grade male students from a primary school in the Netherlands, coined hereafter by the pseudonyms Stuart and Frank, who participated in the lesson with ten other students (aged 8 or 9). Their teacher had eleven years of teaching experience at primary schools. The lesson aimed to foster learning about proportion.

The Mathematics Imagery Trainer for proportions (Trainer) was installed on six touchscreen tablets. Every pair received one tablet and a worksheet. The interface of the Trainer showed two bars that could be moved independently. The colour of the bars changed in response to their position to provide direct feedback. Red indicated the bars were in an incorrect position (not the target proportion); dark green/orange indicated the bars were almost in the right position (close to the target proportion) and bright green indicated the bars were in the correct position, or *green locations* (in line with the target proportion). The bars were in the correct position when they represented a certain proportion. For the proportion 1:2, for example, the right bar should be twice as high as the left bar. Layers with lines and numbers could be added to the Trainers interface.

Three consecutive tasks were included in a case study. In the first task, the students were instructed to find as many green locations as possible within 2 min. The students took turns, one found green locations, and the other counted the number of found green locations, then they switched places for two more minutes (see Fig. 3). In the second task, we asked the students to find a small green location and move the bars upwards while keeping them green. Again students took turns; one tried to keep the bars green while the other student was asked to provide feedback about the others movement. Students were asked to write down a rule that unites all the locations they had found. This instruction aimed to elicit a connection between movement and speech modalities by adding layers with lines and numbers to the Trainers interface. In the third task, students were asked to find a small green location and then move the bars upwards – each student moved one bar.

Instructions given to the teacher included a manual describing the lesson sequence. Every task included three parts: (1) a description of the task, (2) guidance during group learning, including suggestions for feedback on students movement and speech, (3) plenary discussion. The plenary discussion contained suggestions for reflecting on the task, including attention for movement, speech, and the bars visual appearance. Two weeks before the experiment, a researcher met the participating teacher at the school to discuss the lesson and answer the teachers questions.

Data collection and analysis. The learning sequence of Stuart and Frank was video recorded. We transcribed their inferences and attached this transcription to the movie, in the form of synchronized subtitles. In the analysis, we focused on learning along two modalities – *movement* and *speech*. The *end movement-goal* is reaching coordination between hands (Abrahamson & Sánchez-García, 2016), in which both hands would move in a fluent and simultaneous movement, and the interval between fingers/bars will continually grow or shrink while the right bar remains twice as high as the left bar. Accordingly, we crystalized three aspects of *movement*, relying on previous work in the interaction with the Trainer (e.g., Abrahamson, 2015). These aspects of movement were (1) Fluency – Is the movement fluent or discrete? (2) Simultaneity – do both hands move at the same time, or is it one hand at a time? (3) Interval – How the interval between the hands (thus, bars) change? Does it grow or shrink gradually, does it remain constant, or does it change without a clear rule? The *end speech-goal* in the context of Trainer tasks is to express qualitative or quantitative utterances about the proportion between the bars (e.g., twice, half) relations between speeds of movement (e.g., righthand moves faster than the left hand) or the size of the interval (e.g., changing interval) (Duijzer et al., 2017). In our analysis of speech modality, we identified three types of verbal utterances: mathematical words related to the idea of proportions, explications of movement strategies in interacting with the Trainer, and the aspects of movement that comprise them (i.e., fluency, simultaneity, and interval).

We looked at both movement and speech modalities to discern the emergence and development of multimodal voices in the learning sequence by Stuart and Frank – as they worked individually, and in their social interactions with each other, the Trainer, the worksheet, and the teacher. We identify three critical phenomena: a multimodal voice, a multimodal dialogic gap, and a multimodal interaction.

4. The manifestation of multimodal dialogue in group learning of mathematical content

4.1. *Episode A: voices expressed in movement*

A *multimodal voice* in multimodal dialogue is a perspective expressed along various modalities such as speech, hand movement, or eye-gaze patterns (Norris, 2004). Multimodal voices can be manifest in one or some of these modalities. Stuart and Frank's multimodal learning activity started with taking turns in finding as many "green locations" as possible within 2 min (green locations – states of the

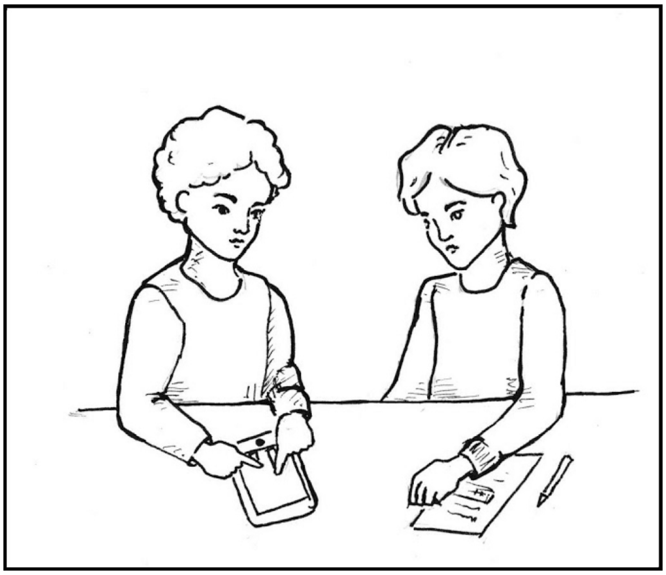

Fig. 3. The initial setting: one student looks for green locations with the Trainer while his peer watches, counts, and records the green locations.

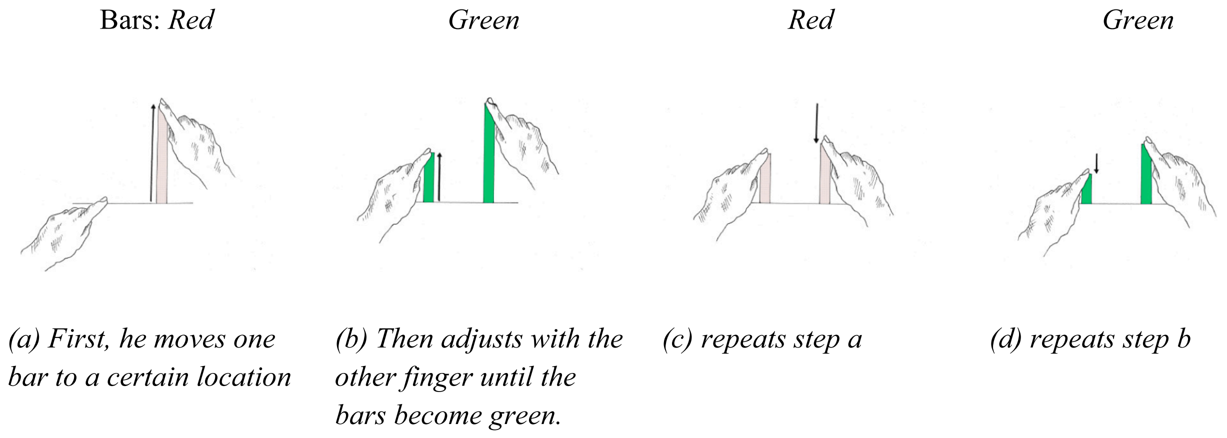

Fig. 4. Stuart's strategy to find green locations.

interface in which the proportion between the bars' heights is 1:2 and the bars are coloured green).

Stuart started looking for green locations, while Frank tallied and recorded these instances (Fig. 3). As a result, Frank and Stuart's multimodal voices were first expressed in movements elicited in the design of the task.

Stuart immediately started finding green locations, by first dragging the right bar to a particular position and then dragging the other bar until both turned green (see Fig. 4). The movement was discrete, non-simultaneous and the interval between the bars changed without any noticeable pattern. Twice in this period, Stuart made attempts to shift his movements to a fluent and simultaneous movement upwards; but in both cases, his hands moved at the same speed, thus keeping a constant interval between the two bars. This strategy was unsuccessful because the bars turned red, and Stuart shifted back to the former strategy. The only speech activation in this sequence was Frank's counting green locations.

In Frank's turn, he developed a different strategy: finding one green location and moving upwards in small steps (Fig. 5). Frank adjusted his index fingers in turns; his right index finger moved a little bit up until the green became darker, then Frank adjusted with the left index finger.

Interactions with Trainers for proportions yield various movement strategies and explanations to these strategies between students. In the time given for each student to interact individually with the Trainer, Both Stuart (Fig. 4) and then Frank (Fig. 5) exhibited non-simultaneous, discrete strategies. However, two distinct voices developed – two movement strategies to find green locations. Stuart exhibited a discrete, non-simultaneous strategy with a changing interval to find green locations. Frank exhibited a gradual strategy, yet again discrete and non-simultaneous. In both of these effective strategies, hand movements were not fluent (the bars did not remain green throughout the movement) and not simultaneous (each hand moves at a time). In Frank's strategy, however, the interval between the bars grew constantly (movement towards *the same* direction). Both applied sporadically a non-productive movement strategy, in which both hands moved simultaneously in a fluent movement but with a fixed interval between them. Both did not follow through as a result of the feedback from the system. Thus, each students interaction with the Trainer elicited idiosyncratic movements to ground two distinct voices, Possibly opening a dialogic gap along the movement modality.

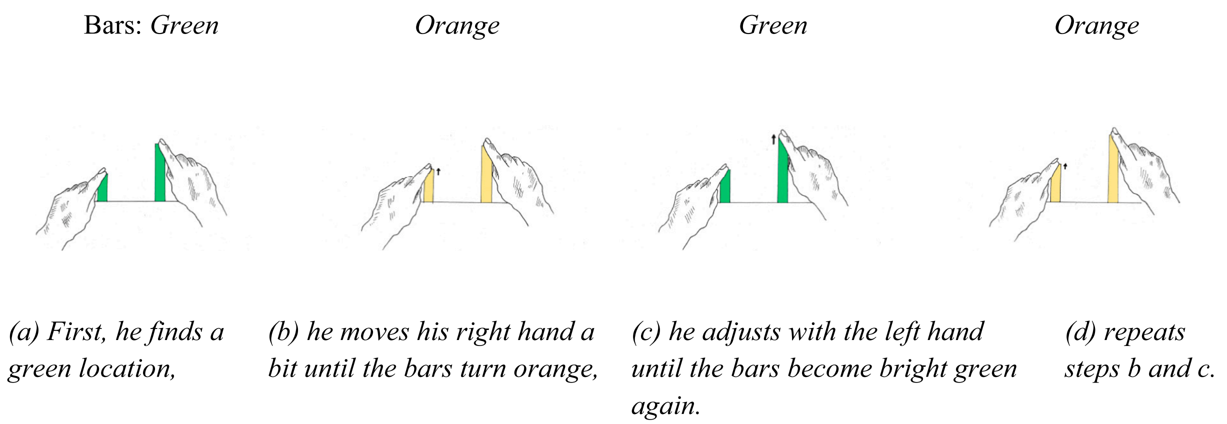

Fig. 5. Frank's strategy to find green locations.

4.2. Minding the movement gap

A *multimodal dialogic gap* is a difference between two distinct multimodal voices, while the gap is within and between modalities. In the following vignette, we describe the first verbal utterances about strategies for solving the task, as Stuart intervenes in Frank's solution. Speech in the transcripts from this point will be highlighted in italics, with brackets where there was a need to support the reader's understanding of the speech. Movement modality will be described in brackets.

A.1 Frank: *This is so easy*. [Leaves the bars in a green location. The right bar takes up two-thirds of the screen.]

A.2 Stuart [points to the right bar]: *That one* [should be] *all the way up*.

A.3 [Frank drags the right bar to the top of the screen. The bars turn red.]

A.4 Stuart [points to the left bar]: *And then that one* [should be] *a bit* [lower].

A.5 [Frank drags the left bar down. The bars are still red.]

A.6 Stuart: *No, up*.

A.7 [Frank drags the left bar upwards until it is at one half of the screen. The bars turn green again.]

A.8 Stuart: *Yes*.

For the remaining time, Frank used a strategy similar to Stuart's initial strategy (Fig. 3).

Movement modality preceded speech modality. As illustrated in Fig. 6, Stuart acknowledged the difference in their movement strategies and attempted to overcome the gap between these two strategies using speech and gesture. He directed Frank's movements to align with Stuart's initial movement [A.2, A.4, and A.6]. Frank listened [A.3, A.5, and A.7] and adopted Stuart's movement strategy. At this point, Frank exhibited two movement strategies for finding green locations. These two movement strategies can be seen as expressions of two distinct voices.

4.3. Minding the multimodal bundle

When the teacher introduced the second task – finding a small green location and moving the bars upwards while keeping them green – she demonstrated this fluent and simultaneous movement in front of the whole class. Then, students were asked to retake turns in dyads: one student practised moving the bars for 3 min, the other student watched and was asked to give feedback about the movement. At this stage, the tablet was still with Frank, and he was the first to solve this task.

B.1 Frank: That is what we did earlier.

B.2 Stuart: Yes, that is easy.

B.3 Frank: Yes, you should not do them the same.

B.4 Stuart: [No], you should do [them the same].

B.5 [Frank puts the two bars on a very small green location and tries to move the bars up with both index fingers.]

B.6 Stuart: They must be green, a small green location.

B.7 [Frank finds a small green location and then moves both index-fingers upwards simultaneously, and the same speed (i.e., fixed interval). The bars turn red.]

B.8 Stuart: They must go the same, and this one [points at the right bar] must always be somewhat higher, right?

B.9 Frank [starts again and multiple times, but never succeeds to drag the bars higher without making them turn red.]: *Ahhh. This is pretty hard*.

Frank started his attempt to solve the task by stating that the strategy for a solution is what they did before [B.1]. Indeed, the teachers strategy resembled the a-to-b strategy that Frank used as his first strategy; however, this time, the task required fluency of movement. Frank [B.7. and B.9] tried to adapt his previous discrete and non-simultaneous strategy into a fluent and simultaneous

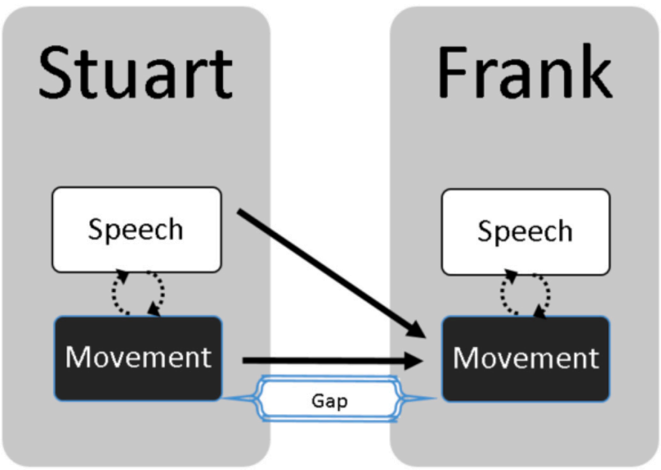

Fig. 6. Stuart identifies a gap between movement strategies and directs Frank's movement using speech and gesture.

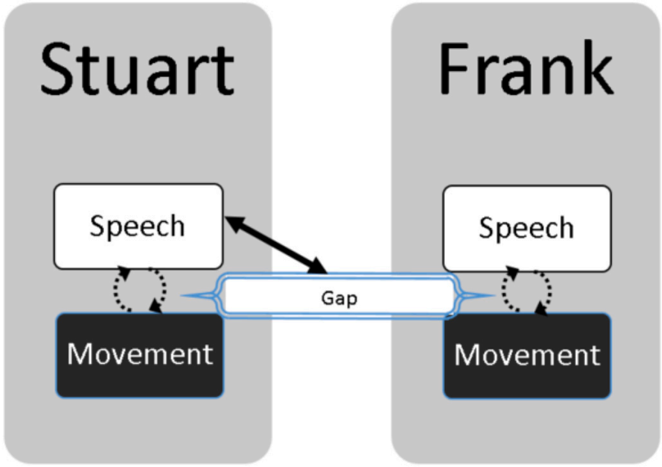

Fig. 7. Stuart identifies a gap between multimodal voices and attempts to reconcile this gap.

movement. Eventually, Frank started over multiple times but could not move his hands fluently and simultaneously while keeping the bars green.

A dialogic gap may be implicit when different voices explicate the same word, and that word can be understood in different discourses (Sfard, 2008). The two students used the term "the same" concerning two distinct attributes of the movement – simultaneity ([B.3] Frank: *Yes, you should not do them the same*) and height ([B.4] Stuart: *Yes you should do* [*them the same*]). Here another dialogic gap became evident. Stuart continued by vocalizing Frank's movements in putting the two bars in a very small green location [B.5, B.6]. Overcoming an incommensurable dialogic gap usually involves regulating discourse on a meta-level (Sfard, 2019), or a voice from a perspective external to the interaction (Wegerif, 2011). After Frank finished the attempt and moved his fingers off the tablet, Stuart addressed the dialogic gap between their two voices: a contradiction in how they related the term "the same" to Frank's movements [B.8]. He does so by explicating two different meanings for this term to Frank [B.8] – for Stuart, *the same* related to simultaneity (*go the same*) and what is not the same are the heights of the two bars (the right bar is always somewhat higher).

This episode exemplifies a multimodal dialogic gap between Stuart and Frank, possibly grounded in the difference between their movement strategies. Frank uses speech modality to comment on interpersonal and intrapersonal in-discrepancies between the voices of the two students. It enabled Stuart to see his thinking from Frank's eyes, refine/re-explicate his thinking about his interaction with the Trainer, and explain this difference to his peer. Later on, in his turn to interact with the Trainer, he would conform to Frank's strategy.

4.4. Interaction within and between modalities

The teacher initiates a classroom conversation to reflect on the experience, in which one student tells the rule "*the right bar should always be bigger than the left bar. And the left bar should always be half as big as the right bar*", which the teacher confirms. This information

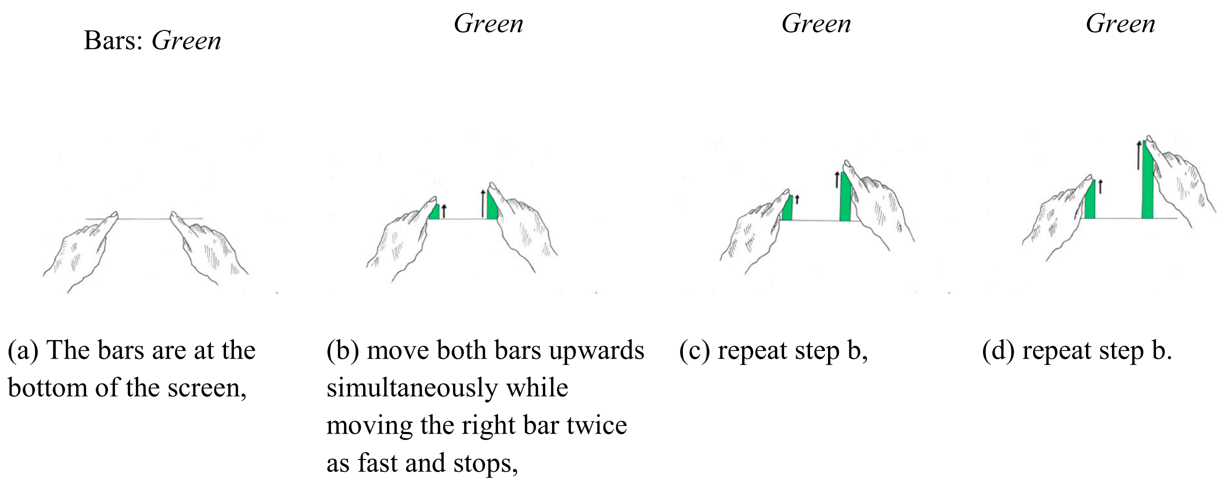

Fig. 8. Stuart and Frank perform a stepwise movement.

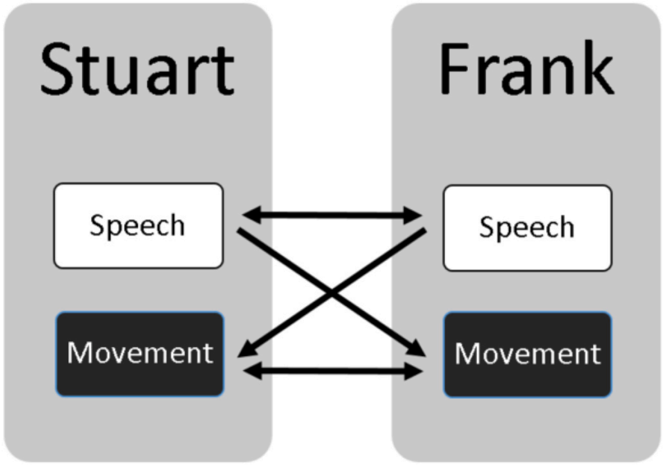

Fig. 9. A model of a multimodal interaction between voices and modalities. Stuart and Frank use speech to coordinate movement and reflect on their joint movement.

helps Stuart to perform a new strategy (while the rest of the students are engaged in a full class discussion): moving the bars simultaneously, in a discrete movement (i.e., not fluent) as the interval grows gradually (see Fig. 8).

For the next task, lines and numbers are added to Trainer interfaces, and students are asked to keep the bars green together. Frank moves the left bar, and Stuart moves the right bar. Stuart choreographs a joint movement where they move together in a discrete and simultaneous movement with an increasing interval between the bars. This joint movement starts as Stuart tells Frank *Now I will go one up and you must be in the middle*, [I go to] *two and* [you go to] *one* and counts down so they can move simultaneously. Later Stuart initiates movement to other green locations. Gradually their movement becomes more synchronized, and less speech is needed to align their actions (Stuart: *now we go to 4*, *so you must go to 2*). Eventually, the students develop a rhythm in which Stuart would name a pair and give a cue when to move *16 and 8*. *3*, *2*, *1*, *GO*! (Fig. 9). Since the interface in this phase also included lines and numbers, the students explicate numerical values. Similar to cases recorded in laboratory settings, the students then reflect on their movements using mathematical frames of reference to explain movement (e.g., Abrahamson & Sánchez-García, 2016). They use speech modality to coordinate and reflect on movement (Fig. 9).

C.1 Frank: We were just taking big steps…. we went from…2. I went from 2 to 4. And you went from 4 to 8.

C.2 Stuart: […] double. [Points to the 8 at the left bar.]. And then from 16 and you to 8.

C.3 Frank: Yes.

C.4 Stuart: And then we went from, uhm. We went… [Moves his hand away from the tablet screen.]

C.5 Frank: Me to 10 and you to 20.

C.6 Stuart: Yes. That.

In the first and second tasks, Stuart and Frank explicated distinct voices. When asked to move together, they manifest a coordinated multimodal interaction with the Trainer task: a simultaneous movement of both bars in the same direction. They use speech to coordinate joint movement in terms of the direction of movement (*Now I will go one up*), respective heights of the bars (*now we go to 4*, *so you must go to 2*), and starting time and duration of the joint movement (*3*, *2*, *1*, *GO*!). This joint movement strategy includes elements of the strategies exhibited by both students: the discrete movement from Stuart's initial strategy and movement of the two fingers in the same direction as in Frank's initial strategy.

After they move together, Stuart and Frank make individual attempts to explain the integration of the two voices to this new voice. Each reflects on the joint solution strategy in terms of location and movement (e.g., *taking big steps*). This reflection upon movement invited the peers to explicate their movement strategies, in congruence with former Trainer studies in which a researcher prompted students to explain their movement strategies (e.g., Abrahamson, 2015; Duijzer et al., 2017). As a result, new mathematical frames of references emerged in the dialogue (*must be at the middle* and *double*).

5. Conclusions

This paper combines dialogic thinking and embodied design theories to create a single, mutually beneficial, analytic lens. We draw on embodied design (Abrahamson, 2015) and dialogic thinking and learning (Wegerif, 2011) to portray three group learning phenomena encapsulated in the multimodal dialogue: multimodal voice, multimodal dialogic gap, and multimodal interaction within and between modalities. We demonstrate these phenomena in the context of the classroom implementation of a lesson aimed to foster tablet-based multimodal learning of the mathematical concept of proportion. The lesson included individual learning before group learning. In the first task, students took turns as one student was asked to find green locations while his peer counted these green locations. We analyze the multimodal – in terms of speech and movement – dialogue between two students, Stuart and Frank.

In Section 4.1, we show how the individual interactions with the Trainer by Stuart and Frank elicited two distinct semiotic bundles

(Arzarello & Sabena, 2014), or multimodal voices. Stuart acknowledged the difference between movement strategies, in Section 4.2. A multimodal dialogic gap requires the existence of (at least) two unique and different voices. In dialogue, people might face incommensurability between how they think of a concept and the concept they attend (Sfard, 2019). In the second task, students took turns as one student was asked to move his hands in a fluent (non-discrete), simultaneous (both hands move at the same time) manner, so the bars remain green; while his peer was asked to comment on this movement. In Section 4.3, we describe Stuart's attempt to settle incommensurability between his and Frank's multimodal voices when they explain the rule they follow in solving the task. Stuart looks at the dialogue from the outside and identifies differences between their semiotic bundles – he referres to the simultaneity of movement as in coordination dynamics (Abrahamson & Sánchez-García, 2016), and Frank referred to the heights of the two bars (Abrahamson et al., 2014). By that, Stuart also brings the two voices together towards closing the multimodal dialogue and convergence to a single coherent voice. Both students individualized this voice, and no multimodal dialogic gap was evident.

How can dialogic pedagogy be used to foster multimodal learning of mathematical content in a mathematics classroom? Fostering dialogue in a lesson is quite challenging since there is an inherent tension between dialogic learning and monologic instructional contexts committed to ideas such as success in curricular activities, correct answers, and teacher controlled discussions (Alexander, 2008; Marjanovic-Shane et al., 2019). This tension is particularly true in mathematics lessons, where success is easily discernable in the form of correct or incorrect answers. These tensions would usually impact teachers attempts to foster dialogic pedagogy, inluencing them to minimize, tame, or extinguish the dialogue in their teaching. Let alone in cases where teachers themselves are not used to teach in a dialogic manner: teachers are often used to control the discourse (Alexander, 2008). In dialogic pedagogy, this kind of control is mostly perceived as authoritative and limiting the dialogue.

A dialogic gap is essential for the dialogue, but nothing guarantees that two students would come up with distinct voices in the interaction with the Trainer (or in other contexts, for that matter). In Section 4.4, Stuart and Frank repeat the second task, but now they controlled one of the two Trainer bars. This element of the lesson design is a relic of former implementations of the Trainer with dyads, in which students were asked to solve the task together from the beginning (Abrahamson et al., 2011). In these cases, the dialogic gap was in speech modality (and maybe eye-gaze) but not in movement. We would argue that eliciting (multimodal) voices before the interaction may increase the chances of a dialogic gap at the interaction's onset. Other practices may include grouping students together in a way that leverages the dialogic gap. Here recent technological developments may help collect data about students learning (Blikstein & Worsley, 2016; Pardos et al., 2018; Xu et al., 2019). Using such analytics can help in an automated grouping of students who explicate distinct voices and make sure there is a dialogic gap at the interaction onset (Abdu et al., 2019). Recently, research on learning with the Trainers has taken first steps into the world of learning analytics, as Pardos et al. (2018) showed their ability to use deep-learning algorithms to classify learners movement strategies with the Trainer. Abdu et al. (2018) used concepts from complex systems theories to create an algorithm that identifies the emergence of steady movement strategies. In the future, we intend to develop tools that identify patterns of multimodal voices and prompt teachers to group learners that exhibit a multimodal gap between them.

How may these findings reflect future learning design? In many cases, an interaction between two novices about a specific topic would not yield a dialogue, since a dialogic gap between the voices merely exists. For example, most humans would agree on the meaning of a term such as "equal" since they have less of a difference in their views on its meaning. In such cases, voices will easily collide, and the dialogue will wither, thus, become a monologue (e.g., Middendorf, 1992). Nevertheless, if we take a linguist and a mathematician to discuss the meaning of the term "equal," they may have different nuanced views on its meaning. If we could summon these imagined linguistic and mathematicians and asked them to understand the differences in their perceptions of the term "equal," there are good chances that the difference between these two voices would have sparked an interaction that stems from an inherent difference in perspectives. Thus, we propose that eliciting voices along several modalities before students interact may increase the chance for a dialogic gap at the interaction's onset.

We propose an analytic lens that combines dialogic learning and embodied design. This analytic lens expands the notion of dialogue as a multimodal activity. Accordingly, we find it imperative to explore how a multimodal dialogue manifests in learning in other contexts within and outside mathematics education. Using these two theories in tandem can also inform small-group embodied learning within a realistic social context such as a classroom. We see the current study as a set point for a productive dialogue between these two theories.

Acknowledgments

This work has been partially funded by the Humanities and Social Sciences Fund of The Israel Academy of Sciences and Humanities and the Netherlands Organisation for Scientific Research (NWO) under grant number 652.001.004.

References

Abdu, R., Abrahamson, D., Bakker, A., & Shayan, S. (2018, October). Applying coordination dynamics to technologically-based embodied mathematics learning. In *Paper presented at the EARLI sig 20 & 26 conference, Jerusalem, Israel*.

Abdu, R., Olsher, S., & Yerushalmy, M. (2019, July). Towards automated grouping: Unraveling mathematics teachers considerations. In B. Barzel et al., (Eds.), Proceedings of the 14th international conference on technology in mathematics teaching – ICTMT Vol. 14, (pp. 147–154). Essen, Germany.

Abdu, R., & Schwarz, B. (2020). Split up, but stay together: Collaboration, cooperation, and mathematical problem solving. *Instructional Science, 48*(3), 313–336.

Abrahamson, D. (2009). Embodied design: Constructing means for constructing meaning. *Educational Studies in Mathematics, 70*(1), 27–47.

Abrahamson, D. (2015). Reinventing learning: A design-research odyssey. *ZDM - Mathematics Education, 47*(6), 1013–1026.

Abrahamson, D., & Abdu, R. (2020). Towards an ecological-dynamics design framework for embodied-interaction conceptual learning: the case of dynamic mathematics environments [Special issue]. In T. J. Kopcha, K. D. Valentine, & C. Ocak (Eds.), *Embodied Cognition and Technology for Learning*. Educational Technology Research and Development.
Abrahamson, D., & Bakker, A. (2016). Making sense of movement in embodied design for mathematics learning. *Cognitive Research: Principles and Implications, 1*(1), 33.
Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM-Mathematics Education, 46*(1), 79–93.
Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences, 25*(2), 203–239.
Abrahamson, D., & Trninic, D. (2015). Bringing forth mathematical concepts: Signifying sensorimotor enactment in fields of promoted action. *ZDM-Mathematics Education, 47*(2), 295–306.
Abrahamson, D., Trninic, D., Gutiérrez, J. F., Huth, J., & Lee, R. G. (2011). Hooks and shifts: A dialectical study of mediated discovery. *Technology, Knowledge, and Learning, 16*, 55–85.
Akkerman, S., & Niessen, T. (2011). Dialogical theories at the boundary. In *Dialogicality in focus: Challenges to theory, method, and application* (pp. 61–72). New York: Nova Science.
Alexander, R. (2008). Culture, dialogue, and learning: Notes on an emerging pedagogy. In N. Mercer, & S. Hodgkinson (Eds.), *Exploring talk in school: A celebration of the work of Douglas Barnes*. London: Sage.
Anderson, M. L., Richardson, M. J., & Chemero, A. (2012). Eroding the boundaries of cognition: Implications of embodiment. *Topics in Cognitive Science, 4*, 717–730.
Arcavi, A., & Isoda, M. (2007). Learning to listen: From historical sources to classroom practice. *Educational Studies in Mathematics, 66*(2), 111–129.
Arzarello, F., Paola, D., Robutti, O., & Sabena, C. (2009). Gestures as semiotic resources in the mathematics classroom. *Educational Studies in Mathematics, 70*(2), 97–109.
Arzarello, F., & Sabena, C. (2014). Introduction to the approach of movement, production, and communication (APC). In *Networking of theories as a research practice in mathematics education* (pp. 31–45). Cham: Springer.
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*.
Bakhtin, M. (1984). Problems of Dostoevskys poetics. Theory and history of literature, Vol. 8. Transl. C Emerson. Minneapolis: University of Minnesota Press.
Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. London, UK: Routledge.
Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3*(2), 220–238.
Bridges, S. M., Hmelo-Silver, C. E., Chan, L. K., Green, J. L., & Saleh, A. (2020). Dialogic intervisualizing in multimodal inquiry. *International Journal of Computer-Supported Collaborative Learning, 15*(3), 283–318.
Buber, M. (1923). *I and thou*. New York, NY: Charles Scribners Sons.
Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1), 15–42.
Cresswell, J., & Teucher, U. (2011). The body and language: MM Bakhtin on ontogenetic development. *New Ideas in Psychology, 29*(2), 106–118.
Duijzer, A. C. G., Shayan, S., Bakker, A., van der Schaaf, M. F., & Abrahamson, D. (2017). Touchscreen tablets: Coordinating action and perception for mathematical cognition. *Frontiers in Psychology, 8*, 1–19.
Flood, V. J. (2018). Multimodal revoicing as an interactional mechanism for connecting scientific and everyday concepts. *Human Development, 61*(3), 145–173.
Flood, V. J., Shvarts, A., & Abrahamson, D. (2020). Teaching with embodied learning technologies for mathematics: Responsive teaching for embodied learning. *ZDM-Mathematics Education, 52*(7), 1307–1331.
Galetzka, C. (2017). The story so far: How embodied cognition advances our understanding of meaning-making. *Frontiers in Psychology, 8*, 1315.
Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596.
Hall, R., & Nemirovsky, R. (2012). Introduction to the special issue: Modalities of body engagement in mathematical activity and learning. *Journal of the Learning Sciences, 21*(2), 207–215.
Hennessy, S. (2011). The role of digital artefacts on the interactive whiteboard in supporting classroom dialogue. *Journal of Computer Assisted Learning, 27*(6), 463–489.
Hetherington, L., Hardman, M., Noakes, J., & Wegerif, R. (2018). Making the case for a material-dialogic approach to science education. *Studies in Science Education, 54*(2), 141–176.
Lee, V. R. (2016). A knowledge analytic comparison of cued primitives when students are explaining predicted and enacted motions. In C. K. Looi, J. L. Polman, U. Cress, & P. Reimann (Eds.)*, Vol. 1. Transforming learning, empowering learners: The international conference of the learning sciences (ICLS) 2016* (pp. 170–177) (Singapore).
Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 445–452.
Marjanovic-Shane, A., Meacham, S., Choi, H. J., Lopez, S., & Matusov, E. (2019). Idea-dying in critical ontological pedagogical dialogue. *Learning, Culture, and Social Interaction, 20*, 68–79.
Matusov, E., Marjanovic-Shane, A., & Gradovski, M. (2019). Bakhtinian pedagogy in conventional educational institutions. In *Dialogic pedagogy and polyphonic research art* (pp. 225–246). New York: Palgrave Macmillan.
Middendorf, M. (1992). Bakhtin and the dialogic writing class. *Journal of Basic Writing, 11*(1), 34–46.
Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. New York: Routledge.
Pardos, Z. A., Hu, C., Meng, P., Neff, M., & Abrahamson, D. (2018). Characterizing learner behavior from high-frequency touchscreen data using recurrent neural networks. In *Adjunct proceedings of the 26th conference on user modeling, adaptation and personalization (UMAP 18)*. Singapore: ACM (6 pages).
Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics, 70*(2), 111–126.
Radford, L., Arzarello, F., Edwards, L., & Sabena, C. (2017). The multimodal material mind: embodiment in mathematics education. In J. Cai (Ed.), *The first compendium for research in mathematics education*. Reston, VA: National Council of Teachers of Mathematics.
Schindler, M., & Bakker, A. (2020). Affective field during collaborative problem posing and problem solving: A case study. *Educational Studies in Mathematics, 105*, 303–324.
Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge University Press.
Sfard, A. (2019). Learning, discursive faultiness, and dialogic engagement. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 89–99).
Shvarts, A., & Abrahamson, D. (2019). Dual-eye-tracking Vygotsky: A microgenetic account of a teaching/learning collaboration in an embodied-interaction technological tutorial for mathematics. *Learning, Culture, and Social Interaction, 22*, Article 100316.
Sinclair, N., Chorney, S., & Rodney, S. (2016). Rhythm in number: Exploring the affective, social, and mathematical dimensions of using TouchCounts. *Mathematics Education Research Journal, 28*(1), 31–51.
Steier, R., Pierroux, P., & Krange, I. (2015). Embodied interpretation: Gesture, social interaction, and meaning-making in a national art museum. *Learning, Culture, and Social Interaction, 7*, 28–42.
Taylor, M. M., Néel, F., & Bouwhuis, D. (2000). *The structure of multimodal dialogue II*. John Benjamins Publishing.
Teo, P. (2019). Teaching for the 21st century: A case for dialogic pedagogy. *Learning, Culture, and Social Interaction, 21*, 170–178.
Wegerif, R. (2006). Dialogic education: What is it and why do we need it? *Education Review, 19*(2).
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.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*(58), 1–13.
Wit, A. P. (2006). Interacting in task groups. In O. Hargie (Ed.), *Handbook of communication skills* (3rd ed., pp. 383–402). London, UK: Routledge.
Xu, F., Wu, L., Thai, K. P., Hsu, C., Wang, W., & Tong, R. (Eds.). (2019). *MUTLA: A large-scale dataset for multimodal teaching and learning analytics. ArXiv preprint arXiv: 1910.06078*.
Yoon, C., Thomas, M. O., & Dreyfus, T. (2011). Grounded blends and mathematical gesture spaces: Developing mathematical understandings via gestures. *Educational Studies in Mathematics, 78*(3), 371–393.
Zhou, Z., Márquez Segura, E., Duval, J., John, M., & Isbister, K. (2019, October). Astaire: A collaborative mixed reality dance game for collocated players. In *Proceedings of the annual symposium on computer-human interaction in play* (pp. 5–18). ACM.
Zohar, A. R., & Levy, S. T. (2019). Attraction vs repulsion–learning about forces and energy in chemical bonding with the ELI-Chem simulation. *Chemistry Education Research and Practice, 20*(4), 667–684.