*Mediterranean Journal for
Research in Mathematics Education
Vol. 17, 25-51, 2020*

Moving Towards Understanding: Students Interpret and Construct Motion Graphs

Carolien Duijzer, Utrecht University, the Netherlands.

Marja van den Heuvel-Panhuizen, Utrecht University, the Netherlands; Nord University, Norway.

Michiel Veldhuis, Utrecht University, the Netherlands; iPabo University of Applied Sciences, Amsterdam, the Netherlands.

Jan Boom, Utrecht University, the Netherlands.

Michiel Doorman, Utrecht University, the Netherlands.

Paul Leseman, Utrecht University, the Netherlands.

Correspondence concerning this article should be addressed to Carolien Duijzer, Freudenthal Group FSW, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands.

Abstract

Bodily experiences are associated with powerful forms of understanding, yet not much research has investigated to what extent bodily experiences benefit the development of graphical reasoning. We examined the effectiveness of providing embodied support in a teaching sequence of six lessons on motion graphs, including both graph interpretation and graph construction activities, on fifth-grade students' reasoning about graphically represented motion. Divided over nine classes 218 students took part in our study. Students in three classes received lessons on graphing motion with direct embodied support, three classes received lessons on graphing motion with indirect embodied support, and three classes served as a baseline condition and received lessons on a different mathematics topic. Development of students' graphical reasoning was measured on four measurement occasions. All students were given these same tasks four times with two months intervals. The teaching sequence on graphing motion took place either after the first, second, or third measurement. We used a cohort-sequential design to assess the intervention effect, the condition effect and the fading effect. Results showed that students improved their graphical reasoning at post-intervention-measurements when compared to their performance before the intervention. Moreover, students in the teaching sequence with direct embodied support showed a slightly larger gain in their graphical reasoning than students in the teaching sequence with indirect embodied support. These results suggest that embodied support as a learning facilitator can improve reasoning about graphing motion in primary school classrooms.

1. Introduction

The ability to understand and reason about graphical representations is a core part of science and mathematics proficiency and, therefore, an important topic in education (OECD, 2000; Roth & Bowen, 2003). Reasoning about graphical representations involves a broad range of skills ranging from encoding basic visual and spatial information in the graph, such as the scaling of the axes, the slope or the intercept, to relating these features to the conceptual or scientific phenomenon they represent, such as a sloped straight line in a distance-time graph reflecting constant speed (Shah & Hoeffner, 2002). Since graphing is often addressed within mathematics lessons, when graphing linear functions, students are mostly confronted with idealized examples, whereas graphs representing real-world phenomena often contain ambiguous elements such as noise or non-linearity (Lai et al., 2016). This might be one of the reasons that students are unable to apply their apparent understanding of graphs within mathematics lessons to graphs they encounter outside the mathematics classroom (McDermott et al., 1987). In the Dutch primary school mathematics curriculum, graphing is only briefly treated. Since graph comprehension – and reasoning about graphs – can be challenging, even for otherwise capable learners and expert users (e.g., McDermott et al., 1987; Roth & Bowen, 2003), it is generally agreed upon that students should be offered ample opportunities to acquire the skills associated with graph interpretation and construction, and to reason about these graphs (e.g., NCTM, 2000; Wang et al., 2012; Wavering, 1989).

In this study, we aimed to foster students' graphical reasoning in primary school. To this end we developed a teaching sequence on motion graphs representing the real-world phenomenon of distance changing over time. In such graphs, students are prompted to connect elements of the graphical representation to the physical event that is represented and to reason about the relationship between the variables on the horizontal and vertical axis as well as their pattern of covariation (Leinhardt et al., 1990). We investigated both short-term and middle-long-term effects of this teaching sequence on students' reasoning about graphs. Following recent proposals to include bodily experiences in teaching graphing (e.g., Duijzer et al., 2019b), stemming from the wider embodied cognition approach to learning and development (see below), we investigated in particular whether a teaching sequence on motion graphs incorporating direct physical experiences has a stronger effect on students' graphical reasoning than a teaching sequence without such direct physical experiences.

2. Theoretical background

2.1. Graphical reasoning

Recognizing visual features of a graph, such as data points and values on the axes, interpreting relationships represented by these features, and connecting these relationships to what the graph actually represents, are three essential processes for comprehending graphs (Shah & Hoeffner, 2002). Graph comprehension is related to developing *graph sense* (Friel et al., 2001; Robutti, 2006). Graph sense, like *number sense* (Resnick, 1989) and *symbol sense* (Arcavi, 1994), is a holistic construct. It is a way of thinking, of becoming sensitive for what various graphs might represent and for how various (non-standard) phenomena might be graphed, both locally and globally. It also includes the ability to distinguish between discrete and continuous representations, to recognize the meaning and significance of the slope, and the more general visual characteristics of the graph (e.g., Robutti, 2006). A student should become flexible in recognizing and using these components, and should also be able to explain their thinking and communicate it to others using graph related language (Friel et al., 2001). When, for example, reasoning about representing the dynamic situation of distance changing over time in graphs, students should be given the opportunity to connect the represented physical situation (i.e., motion) with visual elements of the graphical representation (e.g., the slope, rate of change), and vice versa (e.g., McDermott, 1987). Graph sense encompasses both graph interpretation and graph construction (Friel et al., 2001), although the latter has only rarely been addressed in research on lesson activities (e.g., Leinhardt et al., 1990; Mevarech & Kramarski,1997).

The extent to which students are able to comprehend and reason about graphical representations depends upon many factors such as prior personal experiences, basic everyday intuitions, and familiarity with the graph's conceptual content (Friel et al., 2001; Janvier, 1981; Shah & Hoeffner, 2002; Vitale et al., 2015). When graphs represent changes over time (e.g., increase of distance or length), which are particularly difficult for students to understand (Arzarello & Robutti, 2004), several misconceptions about interpreting and constructing graphs can arise (Glazer, 2011). For example, a student can interpret a graph as an *iconic* representation of a real event (Bell & Janvier, 1981; Leinhardt et al., 1990). This might happen when a student interprets the intersection of two lines in a speed-time graph as the moment when two persons or objects meet. Such reasoning about the graph is not necessarily illogical, because the student simply builds upon informal and intuitive understandings encountered in everyday reality, and applies this knowledge to the graph (e.g., Elby, 2000; Lakoff & Núñez, 2000). Similarly, when asked to construct a graph, a student might draw a line representing the actual path of motion like a map (e.g., McDermott, 1987; Mevarech & Kramarski, 1997). Various studies have shown that such an *iconic* or *pictorial* way of reasoning about graphs representing change over time can be quite persistent in students (Clement, 1985; Mokros & Tinker, 1987). These superficial interpretations might hamper the deeper conceptual understanding of graphs as representing a specific meaningful relationship between more than one variable (Lai et al., 2016; Leinhardt et al., 1990). Being able to resist superficial interpretations and instead draw correct inferences about what a graph actually represents is an important part of graphical reasoning.

2.2. Fostering graphical reasoning

In order for students to develop their graphical reasoning, teachers should preferably build on a students' informal and natural intuitions, and as a consequence circumvent aforementioned misconceptions. It is thus important that students should be offered ample opportunities to discover the deeper relationship between the variables on the axes and reason about their pattern of covariation (e.g., Friel et al., 2001; Lai et al. 2016; Leinhardt et al., 1990; Mokros & Tinker, 1987). Covariational reasoning, for young students, entails the mental coordination of the values of two quantities, while keeping in mind that at every moment the other quantity also has a value (Carlson et al., 2002; Saldanha & Thomspon, 1998). This covariational reasoning is important when interpreting and constructing graphical representations, because it enables students to make a connection between the two variables represented on the graph's axes (Saldanha & Thompson, 1998).

Instructional approaches targeting students' graphical understanding can be divided in two main categories; on the one hand, approaches in which the focus is more on quantitative or local aspects of graphing, on the other hand approaches in which the focus is more on qualitative or global aspects (e.g., Leinhardt et al., 1990). Choosing scales, fitting the paper, reading points in the graph, and letting students plot points from data given in tables, are instructional activities that lead to a focus on graphs' local aspects when interpreting the meaning of a graph or when drawing a graph (e.g., Berg & Smith, 1994; Hattikudur et al., 2012; Lai et al., 2016). When following these more or less fixed routines, a deeper conceptual understanding of the relational aspect of the represented variables might not be sufficiently supported (Yerushalmy & Schwartz, 1993). For example, when a student plots points in a graph and produces a correct slope, this does not necessarily imply understanding of what the slope represents (Vitale et al., 2016). Additionally, Thompson and Carlson (2017) argue how the plotting of points in the graph and "connecting points" without a deeper discussion of the values between successive points, often hampers a deeper understanding of the line in the graph as representing a relationship between two continuously changing quantities.

In contrast, without instructional emphasis on numerals and procedures, students have been found to look at the represented information at a more qualitative and global level (e.g., Krabbendam, 1982). An advantage of a more qualitative, global approach is that it resembles how one might judge a graph in real-life, which often excludes performing calculations on the graph's represented values (Cleveland & McGill, 1984). Another advantage is that when *interpreting* a graph, students can focus on the graph's general shape (Leinhardt et al., 1990), and when *constructing* a graph students can visualize a relationship between two variables as shapes of trends mapped onto the graphs' axes (Matuk et al., 2019). As described by Castillo-Garsow et al. (2013), thinking about the relationship between two variables as continuously changing necessarily involves thinking about motion. This thinking about motion might act as an embodied conceptual metaphor (Lakoff & Núñez, 2000), which maps early everyday experiences with motion to the abstract concept of (graphically represented) continuous change (see also Lakoff, 2014).

In addition to a focus on local or global aspects of graphing, particular learning facilitators that are included in the design of learning environments have been found to foster students' graphical reasoning. For example, in a study by diSessa et al. (1991) students (11-12 years) invented representations of a motion story about a car travelling through the desert by first drawing discrete representations and then moving on to continuous representations of this motion event. This meaningful motion situation and the emphasis on students' own inventions turned out to be powerful learning facilitators for the development of students' qualitative reasoning about these motion representations. Another example can be found in the work of Noble et al. (2004). Sixth-grade students were asked to make block representations of a moving elevator, using physical cubes. The block representations were then transferred into a simulation environment. The elevator in the environment moved in accordance with the motion represented by the blocks. Over the course of the activities, the students were reasoning about the "fastness" of the elevator, without explicitly referring to more quantitative ratio-based descriptions of the movement. Students' reasoning about this particular motion situation was presumed to support more formal reasoning about multiplicative relationships. In both of these examples the real-world context, thus the context of the travelling car and the moving elevator, supported students' (qualitative) reasoning about the (graphical) representations, which allowed them to further develop their formal mathematical reasoning as well as to partake in more conventional graphing practices.

Another often used learning facilitator, already shortly mentioned, known to facilitate students' qualitative reasoning about graphs is the use of real-time motion and simulation environments (Stroup, 2002). For example, in a study of Nemirovsky et al. (1998) students familiarized themselves with the graphical representation of their own movements in front of a motion sensor that was connected to a desktop computer. This approach allowed the students to reason about the relationship between changes in their own movements and the resulting changes in the graphical representation. In learning environments making use of motion sensor technology physical experiences are an explicit part of students' learning activities. Moreover, through the use of motion sensor technology, the line in the graphical representation becomes meaningful to the students since the line in the graph is connected to their own bodily movements andthus in experienced motion (Kaput & Roschelle, 2013). Using motion sensor technology by which a graphical representation appears in real-time also provides a valuable entry-point into reasoning about continuous change represented in graphs (e.g., distance changing over time), because motion experienced with your own body, or observed, must have a value at every point in time. The explicit introduction of bodily experiences in learning activities is in accordance with an embodied cognition approach.

2.3. Enriching graph instruction: An embodied perspective

Learning environments, in which students' own bodily experiences are an explicit part of the learning activities, are also termed embodied learning environments (e.g., Johnson-Glenberg et al., 2014; Skulmowski & Rey, 2018). The ways in which students are provided with opportunities for bodily engagement in learning environments supporting students understanding of graphing motion can vary widely, ranging from whole- or part-bodily movements to observing someone or something else moving (Duijzer et al., 2019b). Including bodily experiences in learning environments is based on the premise that all cognitive processes originate from the perceptions and actions of our body in interaction with our immediate environment (e.g., Pouw et al., 2014; Wilson, 2002). The resulting action-perception schemes are considered to be the fundament of our cognitive architecture. Also, observing movement of others or mentally simulating actions by activating previously acquired action-perception structures are considered to be part of the embodied cognition continuum. Our brain enables us to simulate particular action-perception structures (and invent new ones) (Van Gog et al., 2014), by re-using the sensorimotor circuits of thebrain that were involved in previous experiences of perceiving and acting (e.g., Anderson, 2010; Pulvermüller, 2013). More specific, through the (simulated) enactment of mathematical structures with our body, content-specific action-perception structures evolve which constitute a source-domain that can be metaphorically projected to target concepts (Abrahamson & Bakker, 2016; Lakoff & Núñez, 2000).

In a recent review of research into embodied learning environments (Duijzer et al., 2019b) it was shown that, although physical experiences are often utilized in learning environments supporting students' understanding of graphing motion, not much comparative research into the development of primary school students' understanding of motion graphs has been conducted to date. Of the six studies that did investigate this age group, only one study (Deniz & Dulger, 2012) took a quasi-experimental approach in a classroom setting, the other studies reported (short-term) case studies, involving one or two students (e.g., Ferrara, 2014; Nemirovsky et al., 1998), or observational research (e.g., Anderson & Wall, 2016). Deniz and Dulger (2012) compared two inquiry-based lesson sequences on motion and temperature of which one was enriched with real-time graphing technology and the other with traditional non-digital laboratory equipment. Both lesson sequences incorporated physical experiences, yet only the technology group received immediate feedback provided by the tool. These technology lessons inter alia consisted of specific movements students had to perform in front of a motion sensor (three lessons on motion, three lessons on temperature, six hours in total), which were displayed in real-time on a computer screen. Afterwards the graphs were discussed with the students. Results showed that using the real-time graphing technology significantly improved students ability to interpret motion and temperature graphs.

Based on their systematic review, Duijzer et al. (2019b) concluded that embodied learning environments making use of students' own motion immediately linked to its representation, which was often done through the use of motion sensor technology, were most effective. Thus, embodied learning environments providing students with direct physical experiences have been found to be helpful in supporting students' understanding of motion graphs.

3. The present study

In the present study, we investigated the middle-long-term learning outcomes of a six-lesson teaching sequence, supporting students' reasoning about motion graphs, featuring a particular sequencing of mathematical graphing tasks. Embodied learning environments supporting students' understanding of graphing motion have been found to be effective in small-scale one-to-one settings, however, to date, in the primary grades their effects have rarely been studied in whole-classroom settings (Duijzer et al., 2019b). To this end, we developed a teaching sequence on graphing motion for primary school students. Following the proposal that higher levels of (mathematical) understanding are grounded in physical experiences regarded as embodied cognitions, we developed two parallel versions of this teaching sequence differing in their degree of directness. The teaching sequence in which students were offered direct embodied support, involved graphing activities in which students' own bodily movements were visualized as a line in the graph, using motion sensor technology. The teaching sequence in which students were offered indirect embodied support involved graphing activities that were mostly paper-and-pencil based or projected on the digital blackboard. Students did work with an image of the motion sensor context, but without the presence of the physical tool. A third group of students served as a baseline condition and received lessons on a different mathematics topic.

The study was carried out in primary school classrooms. As a truly randomized design was not feasible, we used a cohort-sequential design with three cohorts which received the lesson sequence in the first, second and third trimester of the school year, respectively. Each cohort comprised of two classes who received either the direct or the indirect embodied support instruction in the trimester where the lesson sequence was provided. A fourth cohort was included as baseline condition. This cohort received a series of lessons on another mathematical topic. We wanted to investigate the potential effects of the embodied learning activities on students' graphical reasoning ability in the context of modelling motion. We formulated the following research question:

To what extent does embodied support in a six-lesson teaching sequence on graphing motion affect the development of students' graphical reasoning?

To assess students' learning progress as a result of the teaching sequence, tests were administered before and after the teaching sequence. The tests consisted of a number of graphical reasoning tasks and required students to explain in writing their reasoning when solving the tasks. Students' written responses were subsequently evaluated with regard to the level of graphical reasoning displayed. We will analyze changes in students' graphical reasoning by performing a longitudinal analysis on the task level following Item Response Theory (IRT), allowing us to model intra- and inter-individual changes in growth. This approach enables us to increase this study's power, to disentangle faulty reasoning from simple mistakes, and to get better insight in changes in levels of reasoning over time. We hypothesize that students taking part in a teaching sequence on graphing motion will, on average, change in their graphical reasoning from lower to higher levels of reasoning more than can be expected based on mere maturation or multiple testing. Additionally, in line with existing research on embodied learning environments, we hypothesize that students receiving a teaching sequence with direct embodied support will outperform students taking part in a teaching sequence with indirect embodied support.

4. Method

4.1. Participants and study design

Schools and classes were chosen based on the willingness of the teachers to participate, resulting in a convenience sample. A total of 237 fifth-grade students from seven elementary schools, divided over nine classes participated in our study. From 19 students we did not obtain written parental consent to collect data. The final sample consisted of 218 students (Grade 5; *M* = 10.47, *SD* = 0.47; 94 female, 43%) divided over two instruction conditions (indirect support condition, *n* = 68; direct support condition, *n* = 70) and a baseline condition (*n* = 80). All schools were located in the area of Utrecht, the Netherlands. The study was conducted between October 2016 and June 2017. The research was approved by the Ethical Review Board of the facultyof Social and Behavioral Sciences at UtrechtUniversity.

All students participated in a teaching sequence of six lessons on graphing motion (with direct and indirect embodied support) or a non-related topic (probability) in the baseline condition as part of their regular classroom instruction. The study adopted a cohort-sequential design, meaning that for each research condition, one cohort of students participated in the teaching sequence in the first trimester of the school year, the second cohort of students in the second trimester, and the third cohort of students in the third trimester. To compose the cohorts, the six classes that would receive the teaching sequence on graphing motion were first clustered in three pairs on matching general school characteristics. Next, in consultation with the teachers, each pair was assigned to one of the three cohorts. Finally, per cohort, the two classes were randomly assigned to one of both instruction conditions. This design allowed us to (1) have the same researcher teaching all the lessons on graphing motion, and (2) to compare the learning curve during the six-lesson teaching with the baseline condition and post intervention conditions (when not yet having had the teaching sequence). Table 1 gives an overview of the study research design.

Table 1.

The Cohort-Sequential Design of the Study

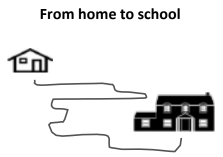
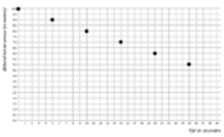
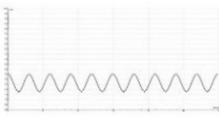
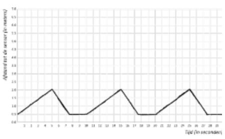
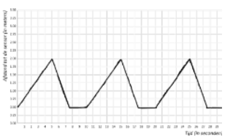
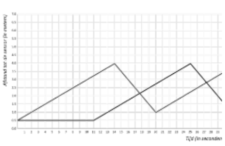
Condition	Cohort		Phase						
				Oct. – Nov. 2016		Jan. – Feb. 2017		Apr. – May 2017	
Baseline	0	(*n* = 80)	M1		M2		M3		M4
With indirect embodied support	1	(*n* = 24)	M1	*Teaching sequence Graphical reasoning*	M2		M3		M4
	2	(*n* = 23)	M1		M2	*Teaching sequence Graphical reasoning*	M3		M4
	3	(*n* = 21)	M1		M2		M3	*Teaching sequence Graphical reasoning*	M4
With direct embodied support	1	(*n* = 21)	M1	*Teaching sequence Graphical reasoning*	M2		M3		M4
	2	(*n* = 22)	M1		M2	*Teaching sequence Graphical reasoning*	M3		M4
	3	(*n* = 27)	M1		M2		M3	*Teaching sequence Graphical reasoning*	M4

4.2. Teaching sequence and procedure

The main goal of the teaching sequence was to help students become acquainted with graphs representing the bivariate relationship of *distance* changing over *time*, and foster students' reasoning about these graphs. The instruction sequence started with informal graphing activities (Lesson 1), followed by a transition from discrete to continuous graphs (Lesson 2), and to continuous graphs (Lesson 3 onwards). Table 2 gives an overview of the teaching sequence, including the main topic per lesson and its key activities.

Table 2.

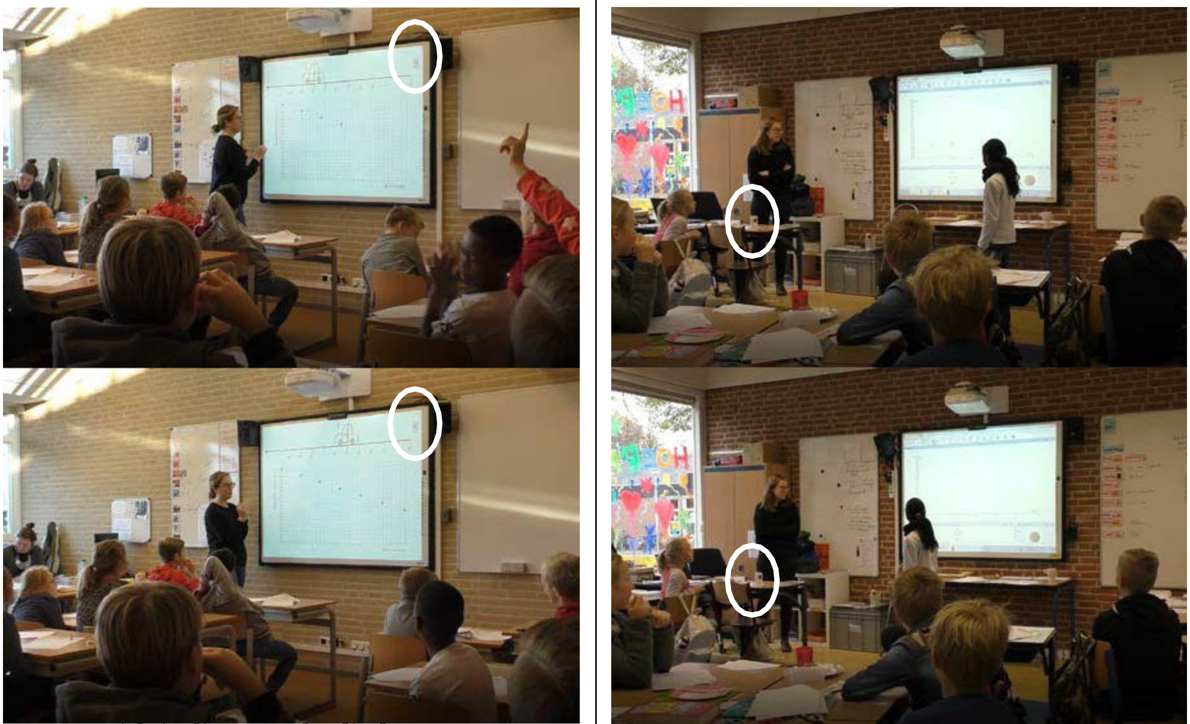
Overview of the Six-Lesson Teaching Sequence on Graphing Motion

Lesson title		Main topic	Key activities
1. Motion: reflecting and representing		Informal graphical representations	*Reason with variables and construct representations of a real-world situation*
2. From discrete to continuous graphs		Measuring distance	*Measure distance in discrete intervals and continuously, and reason about differences between discrete and continuous graphs*
3. Continuous graphs of 'distance to'		Generate, refine, and reason with continuous graphs	*Coupling specific movements to their representation as a line in the graph* *Coupling a concrete situation to a graphical representation*
4. Continuous graphs of 'distance to'		Generate, refine, and reason with continuous graphs	*Coupling specific movements to their representation as a line in the graph* *Investigating how speed is represented in the steepness of slope*
5. Scaling on the graphs' axes		Reason about the relationship between two variables through scaling	*Construct graphs with different scales on the axes*
6. Multiple movements and their graphical representations		Generate, refine, and reason about simultaneous movements and their representation as a graph	*Critically evaluate points of intersection and their meaning*

The teaching sequences in the conditions with indirect and direct embodied support were taught by the first author of this paper, and in the case of direct embodied support with the help of a teaching assistant. Each teaching sequence consisted of six lessons, about 50 minutes each, one lesson per week, divided over 6 weeks. Two weeks before the start of the intervention a general reasoning test was administered. One week before a cohort started with the teaching sequence all students completed the graphical reasoning assessment; this was done for the three cohorts (M1-M3). Finally, after all cohorts had completed the teaching sequence there was a final assessment (M4; see also Table 1).

4.3.1. Instruction conditions

In the instruction condition with indirect embodied support (hereafter: indirect support condition) the students were provided with graphing activities that were paper-and-pencil based (including spoken narratives as well as illustrations of a motion sensor), presented on work sheets or on the digital blackboard. The activities on the digital blackboard were sometimes visualized dynamically, but mostly consisted of non-dynamic illustrations of motion situations of non-human moving objects, such as a toy car travelling a particular distance within a particular period of time. Although the motion situations referred to source domain embodied experiences (e.g., moving your body through space), the graphing activities in the indirect support condition did not involve students enacting the movements in the classroom. Therefore, the degree of embodied support in this instruction condition was low. Similar motion situations and graphing activities were also provided to students in the instruction condition with direct embodied support (hereafter: direct support condition), but instead of only providing the context as an illustrated narrative, students were explicitly prompted to physically enact the situations, using a motion sensor technology. The motion sensor registered enaction and provided students with a direct linkage between their movements and the representation of their movements as a line in the distance-time graph presented on the screen of a computer or the digital blackboard. Therefore, the degree of embodied support in this instruction condition was high. In Figure 1, the difference between both instruction conditions is further explained by giving an example of the lessons' setup. Shown is an activity part of Lesson 2, in which distance is measured at discrete time-intervals (5 seconds).



With indirect embodied support

- Motion sensor illustrated on digital blackboard (see white circle)
- Sensor measures every 5 seconds
- Toy car simulates movement in front of motion sensor
- A graph appears in real-time

With direct embodied support

- Motion sensor physically present (see white circle)
- Sensor measures every 5 seconds
- Student walks in front of motion sensor
- A graph appears in real-time

Figure 1. Difference in set-up between the conditions with indirect and with direct embodied support

4.3.1.1. Motion sensor technology

In the direct support condition we made use of two ultrasonic €Motion sensors, together with Coach6 Software (CMA, Heck et al., 2009). The motion sensor was set to measure the distance between the sensor and the nearest object or person over a 30-second trial, providing a single distance-time graph. The graph was presented on the digital blackboard (Lesson 2 and 6) or on the screen of laptop computers (Lesson 3-5). When moving toward the sensor, the distance between the sensor and the student decreased. When moving from the sensor this distance increased.

4.4. Measures

4.4.1. General mathematics performance

In order to obtain an indication of students' overall mathematics performance, data from the Dutch student monitoring system (CITO LOVS: Janssen et al., 2010), provided by the schools, were used. In this system, schools record their students' results on the biannual standardized mathematics tests. We used the scores of the students on the end-term Grade 4 tests as an indication of their overall mathematics performance (norm population end-term Grade 4: M = 91.9, SD = 10.6, CITO, 2015).

4.4.2. General reasoning

As a measure of students' general reasoning ability, an abbreviated version of the Raven Standard Progressive Matrices (Raven SPM: Raven et al., 2000), consisting of two sets of 9-items, was used (Bilker et al., 2012). Raven's SPM is a test of general reasoning ability and fluid intelligence. Each item consists of a set of pictorial geometric design elements, in black and white. Students are asked to identify the missing element which completes the specific pattern represented by the set. The test was administered to all students in their classrooms during class time, following the instructions in the test's manual.

4.4.3. Graphical reasoning

Students' graphical reasoning about distance-time graphs was assessed four times by a paper- and-pencil test consisting of exactly the same four tasks at each measurement moment: three graph interpretation tasks and one graph construction task. The four tasks were part of a larger test that also included nine other problems related to two other mathematical domains, namely algebra (four tasks) and probability (five tasks). In this study we only include students' performance on the tasks related to graphing motion. Students' received a correctness score on their answer to each task (correct = 1, incorrect = 0; minimum score = 0, maximum score = 4). On Task 2 students could receive partial credit (i.e., resulting in three possible scores for this task "0", "0.5", 1."). In addition, in order to assess students' reasoning, all tasks included an open-ended question, which probed students to make their thinking explicit, by asking them "how do you know?" Students were requested to explain their reasoning in writing and the written responses were coded afterwards for the level of reasoning displayed (see below).

Table 3 shows two tasks as examples. The tasks were developed in such a way that students with different levels of understanding, could show different levels of reasoning in solving them. For example, Task 1 shows a distance-time graph representing the movement of a car.

The speed of the car – the hidden quantity – can be visually deduced by inspecting the steepness of slope. Discovering this hidden quantity can be corroborated with reasoning in which a student explicates that the car in this particular segment travels the largest distance (e.g., when compared to the other segments within the graph), or with reasoning in which a student explicates how the steepness of slope qualitatively represents "distance changing over time" or quantitatively, by taking into account the numerals on the axes. At these higher levels of reasoning a student also reasons about the given quantities on the axes in an (informal) covariational manner.

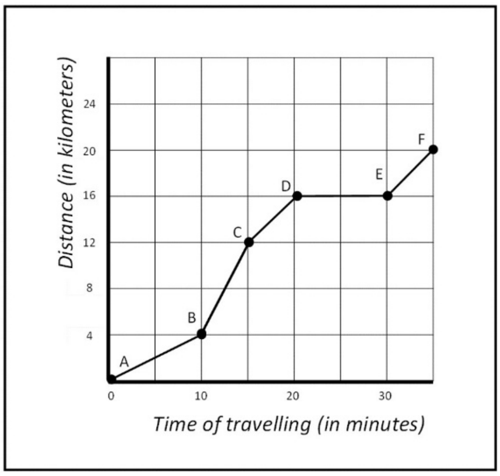
Task 3 represents the graph construction task, including an empty graph and a description of a motion situation. The motion situation consists of three separate parts, in which the train travels at different speeds. Each part of the motion situation implies different rates of change ("twice as fast between 11 and 12 o'clock"). These differences should be made visible by the students in the empty graph. In order to construct a correct graph a student should take into account the relative differences in speed between the three different segments, by quantifying them. In this task, applying the principle "steeper slope means faster movement" does not necessarily result in the correct graph.

Table 3.

Example Tasks Graph Interpretation (left panel) and Graph Construction (right panel)

Task 1

A car drives through town



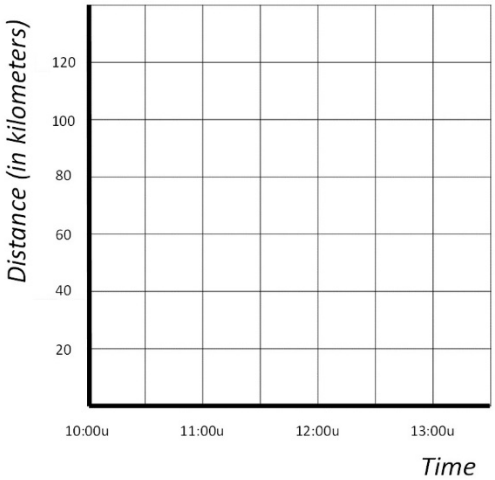
1a. Between which points does the car goes fastest?
1b. How do you know?

Correct answer for this task:
B-C

Score: correct (**1**), incorrect (**0**)

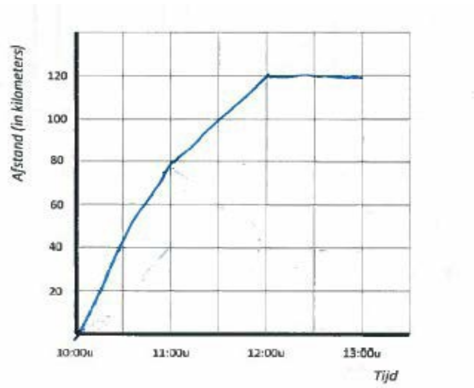
Task 3

A train ride.
A train travels **twice as fast** between **10:00 and 11:00** o'clock than between **11:00 and 12:00** o'clock. The train stands still from **12:00 to 13:00** o'clock.



2a. Draw a graph that fits the description above.
2b. How do you know?

Correct answer for this task:



Score: correct (**1**), incorrect (**0**)

4.4.3.1. Coding scheme for students' level of reasoning

To evaluate students' explanations of how they arrived at a particular solution of the three graph interpretation tasks and the graph construction task, a coding scheme was developed based on an open exploratory analysis of students' explanations. At first, the work of a few students was examined. All research team members first individually categorized these students' responses. Later these classifications were compared, discussed, and revised until agreement was obtained. Finally, this resulted in one coding scheme, applicable to reasoning on both graph interpretation and graph construction tasks, consisting of four categories with increasing sophistication in level of reasoning: *unrelated reasoning (R0)*, *iconic reasoning (R1)*, *single variable reasoning (R2)* and *multiple variable reasoning(R3)*.

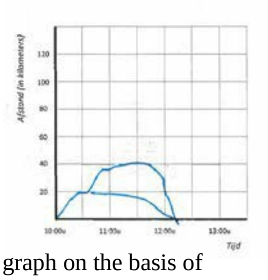
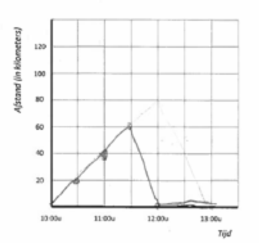
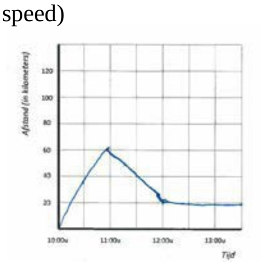
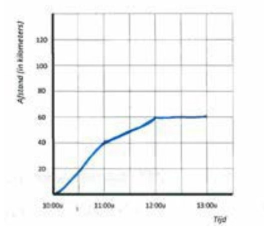
For the graph interpretation tasks, students' written explanations were coded. For the graph construction tasks we took another approach. The students in our sample showed a richness of graphical solutions, yet the majority of the students explained these solutions by simply restating the description of the motion situation as their answer. We assumed students' graphical solutions to be a direct indication of their levels of reasoning outlined above. Therefore, for the graph construction task, we coded students' reasoning as a function of students' ability to correctly take into account the variables on the graph's axes. We distinguished between students who constructed: an illogical graph without taking into account the description of the motion situation (Level R0), a graph based on superficial characteristics of the motion event (Level R1), a graph taking into account a single variable correctly (Level R2), and a graph taking into account multiple variables correctly (Level R3). This highest level of reasoning included, yet was not restricted to, responses that showed a student's informal covariational reasoning.

The coding of the graph interpretation and graph construction tasks resulted in four reasoning scores per measurement moment. In Table 4 the four codes can be found including a description and examples of student's reasoning per category.

An independent second rater coded the four tasks on the four measurements of a subsample of 21 students (336 responses, approximately 10% of all responses). Inter-rater reliability was high with an overall inter-rater reliability of Cohen's Kappa =.92.

Table 4

Coding Scheme used for Students' Level of Reasoning on the Graph Interpretation Tasks (1,2,4) and the Graph Construction Task (3)

		Description	
Level of reasoning	Category	Graph interpretation Task 1 *Example*	Graph construction Task 3 *Example*
		Student reasons…	Student constructs…
Unrelated reasoning	R0	… without referring to the graphical representation or the motion event *"You can see"* *"I guessed"*	… an illogical graph in which the description of the motion event is not taken into account 
Iconic reasoning	R1	… on the basis of the shape of the graphical representation or superficial characteristics of the motion event *"Because those two points are the highest"* *"Over there the line is the longest"*	… graph on the basis of superficial characteristics of the description of the motion event 
Single variable reasoning	R2	… on the basis of a single variable (distance or time or speed) *"Between B and C, the line goes upwards from 4 till 12, so he gives a lot of gas"* *"There he drives 8 kilometers and everywhere else this is 4 or less"*	… graph taking into consideration a single variable (distance or time or speed) 
Multiple variable reasoning	R3	… on the basis of multiple variables (distance and/or time and/or speed) *"The car drives 8 kilometers in 5 minutes. So, in the shortest period of time, the most kilometers."* *"Because, between those two points you find the most kilometers in the shortest time period."*	… graph taking into consideration multiple variables (distance and/or time and/or speed) 

Note. The complete coding scheme, including examples of student responses for each task, can be found in the appendices of this article. Appendix 1 (graph interpretation) and Appendix 2 (graph construction).

4.5. Data analysis

4.5.1. Preliminary analyses and descriptive statistics

We provide sample means and standard deviations for students' general mathematics performance and general reasoning. One-way analyses of variance (ANOVAs) were conducted in order to compare the baseline and the two instruction conditions for differences on general mathematics performance and general reasoning prior to the intervention. A Pearson chi-square test was conducted to test for unintended differences in students' level of reasoning on M1, so before any lessons. Further, we used frequencies of students' level of reasoning (R0, R1, R2, R3) on the graphical reasoning test to calculate the proportion of students using a particular level of reasoning for the baseline condition, and both instruction conditions.

4.5.2. Modelling change in underlyingability

To model students' development in graphical reasoning we adopted an approach in which we combined multi-group Latent variable Growth curve Modelling (LGM), suitable to study longitudinal trends, with assumptions from Item Response Theory (IRT), suitable for categorical data. LGM is a versatile approach for modelling systematic intra- and inter individual differences in change over time and offers many advantages for the modelling of longitudinal data compared to more traditional statistical methods (Willet & Bub, 2005). In our study, we assumed that a student's graphical reasoning would change over the four measurement occasions. We expected a slight increase in reasoning level due to growing familiarization with the tasks and maturation, and a larger increase due to the teaching sequence on graphing motion. The IRT assumption is that graphical reasoning ability itself cannot be directly observed: it is a hypothetical latent ability that underlies the observed reasoning levels in the students' written answers; scored as unrelated reasoning (R0), iconic reasoning (R1), single variable reasoning (R2) and multiple variable reasoning (R3). Thus, the four reasoning levels can be mapped onto the underlying latent graphical reasoning ability. According to IRT, the reasoning levels shown by students on particular tasks are a function of students' unobserved (latent) reasoning abilities and the difficulty of the different levels of reasoning on these tasks. Students' abilities and the tasks' difficulties are placed on the same scale, allowing to express students' reasoning abilities as the probability of showing particular levels of reasoning on these tasks and to express the difficulties of the tasks as the proportions of students showing particular levels of reasoning on these tasks. LGM with IRT yields estimates of students' growth in reasoning ability expressed as the increased probability of showing higher levels of reasoning on a particular set of tasks.

We estimated students' individual growth trajectories based on four partial individual effects. Students may show individual differences in their reasoning on the pre-measurement (*intercept effect*) and in the rate of change over time (*slope effect*) for the subsequent three measurements. In addition to the *intercept* and the *slope effect*, we included an *intervention effect* and a *weakening effect*. With the *intervention* effect we model students' change in ability after partaking in the teaching sequence. For example, an intervention between M1 and M2, might lead to a change in students' graphical reasoning ability between the measurements on M1and M2, and may extend to a change between M3 and M4. The *weakening* effect takes into account the possibility that the intervention effect might fade-out over time.

Two control variables (general mathematics performance and general reasoning) were included as predictors in the LGM analyses to control for individual differences in general mathematical ability and general reasoning ability. Finally, to answer the main question of the current study, condition was added as a predictor into the model since we assumed that the intervention effect might depend on the specific condition students are in (indirect or direct embodied support). Hence, by adding condition as a predictor we could investigate whether the instruction condition impacted changes in students' reasoning ability over time, thus answering the question whether students in different instruction conditions differ in growth trajectories. In a stepwise procedure we first estimated an unconditional model that served as our baseline model only including the intercept effect and the slope effect. In the next step we added the intervention effect and the weakening effect. We then added the two general measures (general mathematics performance and general reasoning) as predictors of the intercept and the slope effect. Both predictor variables were grand mean centered. In the final step, we added condition as a predictor of the intervention effect.

The multi-group latent growth curve model, with time varying effects added, was estimated using Mplus (Version 8; Muthén & Muthén, 2012-2017). A logit link was used to map the likelihood of using a certain level of reasoning (Level R0, R1, R2, or R3) onto students' latent graphical reasoning ability. The logit link implies that we had to use robust Maximum Likelihood Estimation (MLR). As a consequence, because MLR provides no chi-square goodness of fit index, we used the Aikake Information Criterion (AIC) and the Bayesian Information Criterion (BIC) as relative overall fit measures. We report the change in AIC (ΔAIC) and BIC (ΔBIC) for each comparison between models. Both fit indices take into account sample size and the number of parameters. We followed the commonly applied rule that lowest AIC and BIC represent the best model fit. Further, we provide parameter estimates and significance values of the separate effects and the predictors.

4.5.3. Missing data

Of the 218 students in this study, 213 had complete data on general mathematics performance, and 217 had complete data on general reasoning. For the students with missing data on these measures, values were imputed based on class averages. Four students in the conditions with direct or indirect embodied support missed either M2 or M3, while the subsequent measure was present. To avoid having missing post-measurements, we decided to substitute the missing measurement point with the subsequent one. For example, a student in Cohort 1, receiving the intervention between M1 and M2, missed M2. For this student we treated M3 as if it were M2 and M4 as if it were M3.

5. Results

5.1. Preliminary analyses and descriptive statistics

There were no significant differences between the baseline condition and the two instruction conditions on students' general mathematics performance ($F(2, 210) = 0.77$, $p = .465$, partial $\eta^2 = .007$), general reasoning ($F(2, 214) = 0.29$, $p = .752$, partial $\eta^2 = .003$), and level of graphical reasoning on M1 ($\chi^2(6) = 10.88$, $p = .092$). Table 5 presents per condition, for each cohort, the means and standard deviations of general mathematics performance and general reasoning, as well as the correctness scores on the graphical reasoning test for all four measurement moments.

Although they did not have an intervention on graphing motion, students in the baseline condition did seem to improve in their correctness scores over the school year (+ 0.70), as did students in the indirect (+1.03) and direct support condition (+ 1.08).

Table 5

Means and Standard Deviations for General Mathematics Performance, General Reasoning (Abbreviated Raven's SPM) and Graphical Reasoning $M_{measure\ 1} – M_{measure\ 4}$ (correctness scores)

		General measures		Macro measures			
		Mathematics performance CITO E6	General reasoning Raven's SPM	M1	M2	M3	M4
	Cohort	*M*(*SD*)	*M*(*SD*)	*M*(*SD*)	*M*(*SD*)	*M*(*SD*)	*M*(*SD*)
Indirect embodied support [Instruction Condition 1]	1	96.75(9.07)	9.30(2.74)	**1.25(0.97)**	**2.46(0.96)**	2.38(0.86)	2.57(1.08)
	2	100.38(8.16)	12.23(2.81)	2.27(1.13)	**2.93(1.02)**	**2.83(1.02)**	3.20(0.75)
	3	88.67(12.84)	10.04(2.72)	1.98(1.27)	2.36(1.37)	**2.50(1.21)**	**2.75(0.85)**
Total		97.88(10.16)	10.52(2.97)	1.81(1.19)	2.59(1.13)	2.56(1.03)	2.84(0.93
Direct embodied support [Instruction Condition 2]	1	89.75(13.68)	10.50(2.47)	**0.93(0.78)**	**2.35(0.85)**	2.05(1.01)	2.31(1.30)
	2	101.73(12.16)	11.65(2.79)	1.34(0.96)	**2.16(1.10)**	**2.86(0.98)**	2.50(1.12)
	3	94.48(11.63)	10.24(2.43)	2.14(1.31)	2.60(1.42)	**2.26(1.40)**	**2.90(1.23)**
Total		95.42(13.13)	10.81(2.61)	1.52(1.17)	2.38(1.17)	2.39(1.22).	2.60(1.23)
Baseline Condition	4	97.31(12.77)	10.49(2.74)	1.45(1.06	1.91(1.18)	1.97(1.23)	2.15(1.20)

Note. E6=End-term Grade 4. Boldened scores indicate scores on measurements in between which the intervention took place.

The development of students' level of reasoning on the graphical reasoning test for all four tasks together is shown in Figure 2 for the baseline condition. Proportions of students using a particular level of reasoning are shown for each measurement occasion. Students showed some decline of R1 reasoning over time, but a slight increase of R0 and R3 reasoning. Overall, students' level of reasoning (R0-R3) in the baseline condition stayed rather stable overtime.

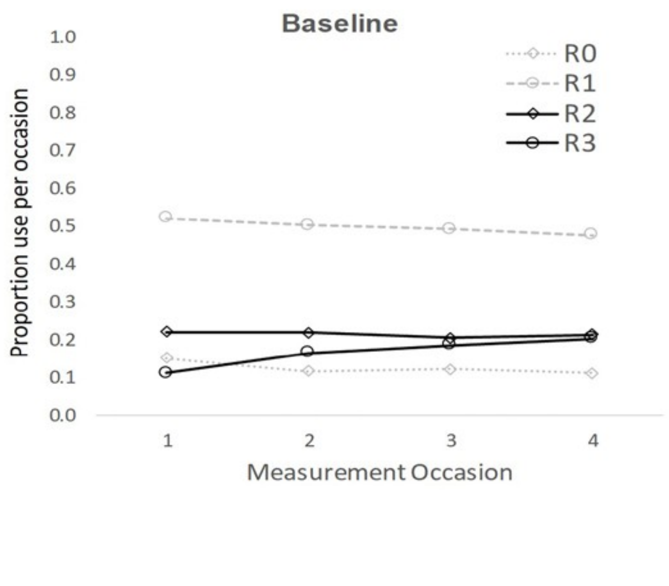


Figure 2. Proportion of students using a particular level of reasoning (R0-R3) for each measurement occasion for the Baseline condition.

In Figure 3, the development of students' reasoning is shown for the indirect support condition (left panel) and the direct support condition (right panel). In this figure measurement occasions are aligned between cohorts, such that the intervention is set to start and end at the same virtual time points for all cohorts. This alignment was necessary in order to be able to visually compare the development of students in the different cohorts, since students in the different cohorts participated in the teaching sequence in different time periods. Students in Cohort 1 participated in the teaching sequence during the first time-period (October – November), directly after the first measurement occasion; students in Cohort 2 received the teaching sequence during the second time-period (January – February), performing two measurements before the teaching sequence; and students in Cohort 3 participated in the third time-period (April – May), performing three measurements before the teaching sequence. When aligned in Figure 3, students in Cohort 1 are shown as having participated in virtual measurements 3 to 6, students in Cohort 2 in virtual measurements 2 to 5, and students in Cohort 3 in measurements 1 to 4. This allows for a direct comparison of the improvement of students in all cohorts following their participation in the teaching sequence by inspecting the change between virtual measurement occasions 3 and 4.

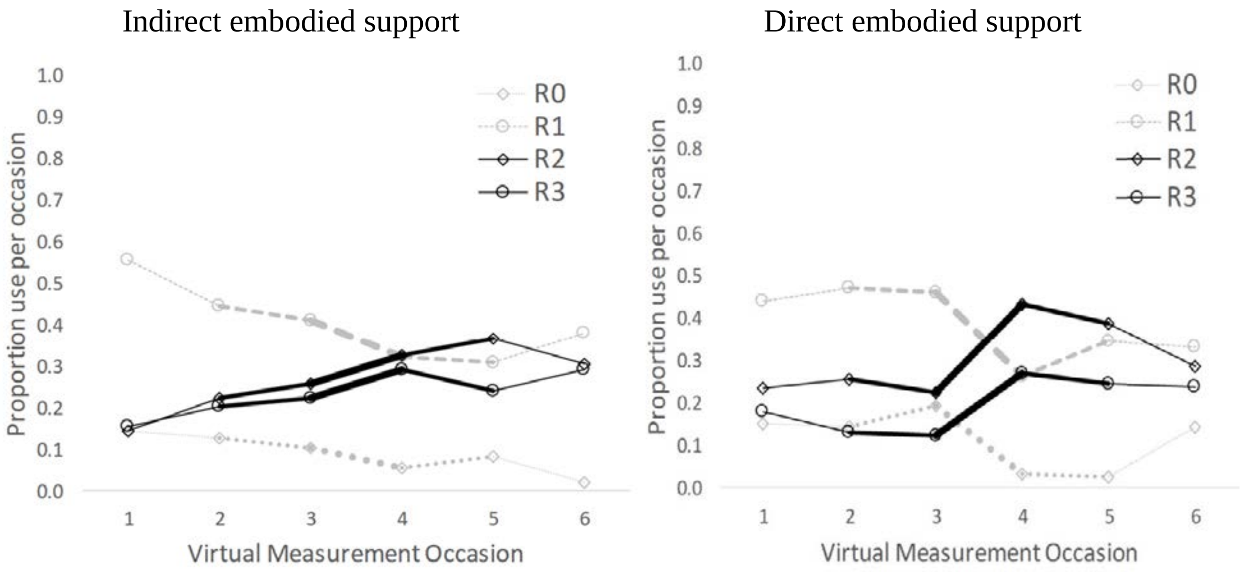


Figure 3. Proportion of students using a particular level of reasoning (R0-R3) for each measurement occasion. Between virtual measurements 3 and 4 the teaching sequence took place for the indirect and direct support conditions. Thin line-segments are based on one cohort, thicker line-segments on two cohorts, and thickest line-segments are based on all three cohorts.

After partaking in the teaching sequence more students in both the direct and indirect support condition showed reasoning on the basis of a single variable (R2) as well as reasoning on the basis of multiple variables (R3).

Additionally, students in the direct support condition exhibited a larger gain in the frequency of R2 and R3 reasoning (R2 + 21% points and R3 + 13% points) than students in the indirect support condition (R2 + 7% points and R3 + 9% points).

5.2. Effects of embodied support on students' graphical reasoning ability

To investigate the general effectiveness of both instruction conditions in terms of immediate (post-test) and middle-long-term (follow-up) effects, latent growth curve analysis was used to model intra-individual change in graphical reasoning over the four measurement points, corrected for general mathematics ability and general reasoning. First, an unconditional growth model, including the intercept effect and the slope, but no other effects was estimated. The fit of this model (AIC = 7970.16; BIC = 8031.08) serves as our baseline. Adding the intervention effect and the weakening effect to the model resulted in an improvement in the overall relative model fit ($\Delta AIC = 83.69$; $\Delta BIC = 73.54$). In addition to the overall fit measures also structural parameters of the model are of interest (Wald tests). The effect of the intervention on students' reasoning was significant (1.10, $p < .001$). There was also a significant weakening effect on the delayed measures after the intervention (-0.47, $p < .001$). The addition of general mathematics performance and general reasoning as predictors of the intercept further improved our model ($\Delta AIC = 86.21$; $\Delta BIC = 79.44$). Both predictors are significant predictors of the intercept effect (general mathematics performance: 0.52, $p < .001$, general reasoning: 0.23, $p = .001$).

To investigate the effect of embodied support on students' reasoning about motion graphs on the immediate and delayed post-test, instruction condition was added as a predictor of the intervention effect. In this way we modelled the relationship between students' changes in graphical reasoning over the four measurement points and the specific condition they are in.

After adding the condition effect to our model, we found an improvement in model fit ($\Delta AIC = 7.64$; $\Delta BIC = 4.25$). Condition turned out to be a significant predictor of the intervention effect ($p = .001$), explaining 25% of the variance of the intervention effect. Thus, students receiving direct embodied support during the teaching sequence displayed higher levels of reasoning after the intervention than students that received indirect embodied support.

In order to gauge the effect of instruction condition, it is helpful to visualize the results. Figure 4 shows these effects for the baseline (left) and the three cohorts separately. The lines in the graphs show the visualization of the additive relationship between the intercept effect, the slope effect, the intervention effect, and the weakening effect, for students in the direct support condition (top line) and students in the indirect support condition (bottom line).

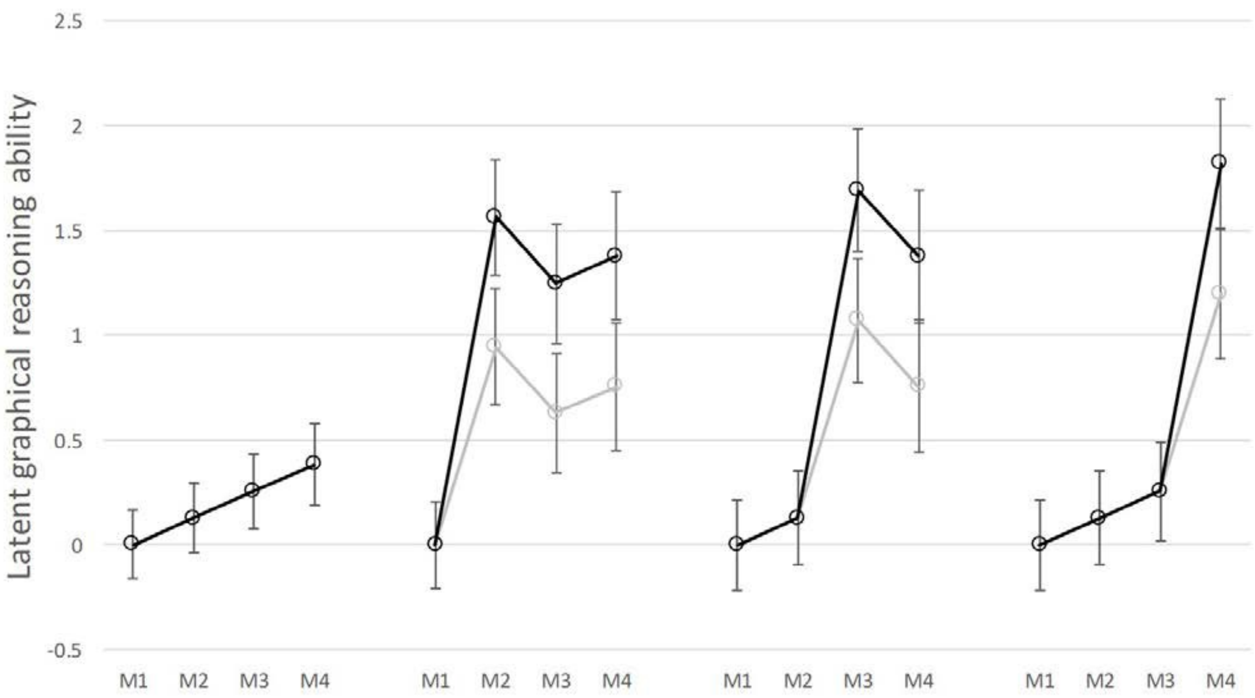


Figure 4. Additive relationship between the effects of the intercept, the slope, the intervention, and the weakening separately for the baseline (left), Cohort 1, Cohort 2, and Cohort 3, for the direct support condition (black line) and the indirect support condition (grey line), for students' latent graphical reasoning ability: Error bars indicate 95% confidence intervals.

Table 6 presents the fit indices and parameter estimates of our final model including all four partial effects (i.e., intercept, slope, intervention, weakening), as well as the three predictors (general mathematics performance, general reasoning, condition).

Table 6.

Fit Indices and Parameter Estimates of the final LGM Model Including all Partial Effects, Control Measures, and the Effect of Condition

Model	AIC/BIC	df	Model parameter	Estimate	*p*-value (two-tailed)
Intercept, Slope,	7792.624/	193	Intercept (mean)	0.0	fixed
Intervention effect,	7873.852		Slope (mean)	0.128	.003
Weakening effect +			Intervention (mean)	1.125	< .001
Condition as predictor of the			Weaken (mean)	-0.443	< .001
intervention effect)			General mathematics performance (mean)	0.044	< .001
+ General mathematics performance			General reasoning (mean)	0.088	.001
+ Non-verbal reasoning			Condition (regression β)	0.309[1]	.001

Note. [1]Condition was coded as 1 Direct support condition and -1 Indirect support instruction condition

5.3. Reaching higher levels of reasoning: Example of two growth trajectories

In order to explicate what the above quantitative analysis implies in relation to the activities that were conducted in the classroom, and the reasoning of the students on the tasks used to assess their levels of reasoning, in this final section we provide the growth trajectories of two students over the schoolyear (see Table 7). We focus on Task 1. The trajectories given below are not representative for the entire sample of students, they serve as an illustration. Both trajectories show growth in reasoning ability as a result of the intervention and some post-intervention fading of this effect. Following the findings of the quantitative analysis, indicating that the direct support condition was more effective on students' growth in graphical reasoning, we restrict ourselves to the instruction condition offering direct embodied support.

Table 7

Growth trajectories of Elliot and Levi showing their reasoning on the four measurement moments

Name	M1	M2	M3	M4
Elliot	[CD-EF] *"I think so, because these are small pieces"*	[B and C] *"because in 5 minutes they travel 12 kilometers"*	[BC] *"Because between these points you have the most kilometers in a short time period"*	[BC] *"I looked and then I have written down the answer"*
Levi	[B-C] *"It is the longest"*	[b and c] *"I looked at which one was the longest and the time"*	[b to c] *"I looked at where the lines in the graph were going up the highest"*	[BC] *"Nowhere it goes as fast in the graph. He travels in 5 minutes, 8 kilometers, he never does this at another moment in the graph"*

Trajectory 1 – Cohort 1: Elliot

On the measurement before the intervention (M1), Elliot based his answer on some superficial characteristics of the graph, resulting in Level R1. The answer of Elliot is "C-D and E-F", which is an incorrect answer. Elliot corroborates his answer with: "Because these are the shortest pieces". With shortest pieces this student refers to the line segments in the graph. On the measurement directly after the intervention (M2) the reasoning of Elliot has changed. He now uses the variables distance and time in an informal covariational manner: "because in 5 minutes they travel 12 kilometers", using both quantities represented on the axes of the graph in his reasoning. On the third measurement moment (M3), Elliot still reasons according to the highest level (R3), still showing reasoning in an informal covariational manner, yet without explicitly mentioning the numerals. Instead he qualitatively refers to the given quantities "most kilometers" and "little time". On the final measurement (M4), Elliot does not show reasoning that is related to the graphical representation anymore. Instead his reasoning is merely procedural, resulting in Level R0. The growth trajectory of Elliot illustrates how a student can show an increase in level of reasoning from pre- to post intervention and a weakening effect on one of the delayed measures, as was found in the quantitative analysis described above.

Trajectory 2 – Cohort 3: Levi

On the first measurement moment, Levi shows reasoning according to Level R1, see Table 7. He, like Elliot, focuses on a particular line segment being "the longest". Although his answer is correct: "BC", the reasoning associated with his answer can be considered superficial. On the second and third measurement moment, without having had an intervention, Levi shows reasoning according to Level R2. For example, on measurement moment 2 he states: "I looked at which one was the longest and the time". Although the first part of this answer is similar to his answer given on measurement moment 1, this time he corroborates his answer with explicitly mentioning the variable time, indicating that he incorporated the quantity time given on the *y*-axis of the graph. Finally, on the fourth measurement moment, directly after having had the intervention, he shows reasoning according to Level R3: "Nowhere it goes as fast in the graph, he travels in 5 minutes, 8 kilometers, he never does this at another moment in the graph." The growth trajectory of Levi shows how Levi throughout the schoolyear shows growth, regardless of having had an intervention. Yet, his reasoning after the intervention clearly is more elaborate.

6. Discussion

In this study, we examined whether a six-lesson teaching sequence on motion graphs raised students graphical reasoning. We defined graphical reasoning as a mixture of qualitative and quantitative reasoning about a single variable or about multiple variables, as opposed to reasoning in an iconic or pictorial way. We took students' written responses to the open-ended graph interpretation and graph construction tasks as reflecting their reasoning and coded this reasoning on four levels of increasing complexity and appropriateness. In line with previous research, the present study investigated the added benefit of direct bodily experiences, compared to indirect bodily experiences in the teaching sequence. We thus asked: *To what extent does embodied support in a teaching sequence on graphing motion affect the development of students' graphical reasoning?* The teaching sequence focused on problem situations involving motion, situated in a real-world context that was presented on worksheets and modelled on the digital blackboard in the instruction condition offering indirect embodied support and was presented on paper and physically enacted in the instruction condition offering direct embodied support. In our method and analyses, we took into account both short-term and long-term effects of the intervention.

We modelled individual changes in graphical reasoning ability using latent growth modelling. We found that students' graphical reasoning improved after taking part in the teaching sequence on motion graphs. Students more often used reasoning taking into account a single variable (Level R2) or taking into account multiple variables (Level R3).

We also found that students taking part in the direct embodied support condition benefited more from the intervention than students in the indirect embodied support condition. Students receiving direct embodied support showed more often higher levels of graphical reasoning (Level R2 and Level R3) after partaking in the teaching sequence than students receiving indirect embodied support. This shows that an embodied learning environment incorporating immediate whole-bodily motion activities is more helpful in stimulating students' reasoning about graphs than when students do not perform immediate whole-bodily motion activities, and instead receive an illustrated model of this motion sensor context on worksheets and the digital blackboard. This finding underscores previous research within this specific mathematics domain (e.g., Deniz & Dulger, 2012), and other mathematics domains (e.g., Fisher et al., 2011). For a review on this topic, see Duijzer et al. (2019b). The difference in terms of estimated abilities, between the two conditions, was about one standard deviation apart. The proportion explained variance, however, was small ($r^2 = .25$). This can be explained by the fact that students in the indirect support condition, were also confronted with activities that capitalize on bodily-based experiences. For example, the object of the toy car used in the indirect support instruction condition, to some extent, might have caused neural activity in the human brain similar to the neural activity induced when viewing another person's action or performing an action (see also Beauchamp & Martin, 2007; Chao & Martin, 2000; Chouinard & Goodale, 2010). Additionally, the graphing of motion itself capitalizes on experienced motion, whereby these experiences with real motion can act as metaphorical mappings between source domain experiences (such as real movements through space) and the graphical representation, even in the absence of direct physical experiences (e.g., Barsalou, 1999; see also Castillo-Garsow et al., 2013).

In previous research it has been established that when students partake in graphing activities, using for example a motion sensor and desktop laptop, several graph reading errors, such as iconic and pictorial interpretations of graphs can be overcome (e.g., Brasell, 1987; Deniz & Dulger, 2012; Duijzer et al., 2019a, Mokros & Tinker, 1987). These findings were mostly based on tests consisting of multiple-choice questions. In our study, we added complexity and depth to the analyses by taking into account students' written explanations as indications of their level of reasoning and changes therein over a prolonged period of time. We illustrated these changes by incorporating two qualitative examples presenting the growth trajectories of two students. At the highest level of reasoning (Level R3) these students reasoned about the variables distance and time in an informal covariational manner. Additionally, these qualitative examples showed the added value of including students' written explanations in the statistical analysis. For example, Levi gave the correct answer on each of the four measurement moments, yet his written explanation show a clear increase in the level of understanding over time. At the first measurement, he incorporates a superficial characteristic of the graph in his reasoning, while at the final measurement (M4) his reasoning changed to reasoning in which he took into account both variables. Thus, including students' written explanations gave us more information regarding their understanding than when we would have only looked at students' correctness scores. This approach is in line with Lai et al. (2016), who show the importance of incorporating a direct measure of reasoning by giving students the opportunity to elaborate on their answers in achievement tests. In this sense, we demonstrated that students' reasoning taking into account iconic or pictorial aspects of the graphs (Level R1), was often replaced by reasoning in which they took into account one or more of the relevant variables (Level R2 and Level R3), regardless of the correctness of their answer.

6.1. The value of direct versus indirect embodied support

The motion sensor context used in our study is just one example of digital technology that has been utilized over the past couple of decades to support learning in mathematics and science classrooms. The digital element of the motion sensor entails the real-time translation of movement into a digitalized graphical representation of that movement. The context of the motion sensor was used extensively in the teaching sequence offering direct embodied support. In the instruction condition offering indirect embodied support, the students did not have the opportunity to benefit from a motion sensor in the physical way. They were offered this context on paper and on the digital blackboard. Thus, on the basis of our comparison between instruction conditions, we cannot determine exactly which specific elements of the teaching sequence were most helpful in facilitating student's graphical reasoning. Both instruction conditions involved sense making activities that were perceptually experienced (Barsalou, 1999; see also Goldman, 2012).

Further, we operationalized direct embodied support as making whole bodily movements in front of the motion sensor. Yet, due to the nature of the motion sensor context, the whole bodily motion activities in front of the sensor to some extent has more advantages than the physical experience of motion alone. It includes physical movement *as well as* immediate feedback provided by the tool. Even though this immediate feedback was sometimes also provided to the students in the teaching sequence with indirect embodied support, the combination of physical experiences with real-time feedback in one instruction condition makes it difficult to disentangle their respective unique effects. Future research could address this by creating a condition in which students for example do not receive immediate real-time feedback, but delayed feedback (see also Brasell, 1987), to isolate the effects of the real-time feedback provided by the tool. Another possibility is to isolate the unique contribution of own bodily motion experiences. For example, by letting students work with a dynamic model of the activities' set-up. An example of such a learning environment is presented in the study of Salinas et al (2016), who gave students the opportunity to control an animated avatar in a computer software program. The movement of the avatar is presented alongside the corresponding graph. The students could influence or control the motion of the avatar, but could not move their selves, eliminating the possibility of direct physical experiences.

6.2. Limitations, strengths and future research

`This study has some limitations that we have to mention here. First, even though students' reasoning on the test items provided us with a window into their thinking processes, we cannot be sure that we captured the full breadth of students' understanding, when *only* looking at their written responses to the tasks. It might be worthwhile to include more extended measures such as think-aloud protocols when solving the tasks. A second, related limitation, is that we included only four tasks to measure students' development in reasoning about motion graphs. Even though using few tasks is a considerable advantage when thinking about the mental effort imposed on the students, future research might consider using more tasks, specifically more graph construction tasks. A third, and final, limitation worth mentioning is that even though we have investigated the teaching sequence in a realistic classroom setting, which enhanced the ecological validity of our study and the applicability of the approach in education, a drawback of this approach is that some of the teaching time was consumed by the procedural aspects of setting up the equipment.

Also, the use of motion sensor technology in the classroom might have had a distracting effect as well. Since not all students are walking at the same time in front of the sensor some students sometimes were disengaged, either by the other small group working with the sensor, or by talking with their peers (see also Anderson & Wall, 2016). A suggestion for future research is to let students work in even smaller groups (e.g., three or four students) on the tasks.

This study also has several strengths. First, an important difference between previous research on graph understanding in the primary grades and the current study is that we looked at the development of students' graphical reasoning over a year. We included multiple measurements to look at students' longitudinal development and to take into account fade-out effects of the intervention. We indeed found a fade out effect for the intervention. Second, from a statistical point of view, this study is innovative in the sense that the used latent growth curve model incorporated categorical responses to the tasks, which allowed us to model gradual changes in levels of reasoning (Boom & Ter Laak, 2007). Third, our cohort-sequential research design enabled us to "re-use" student groups per instruction condition, whereby the groups served as their own control group, depending on the specific cohort. This resulted in the need of fewer participants overall, which is an advantage from both a practical and ethical point of view. Fourth, we incorporated a baseline condition that helped us to more accurately estimate the intercept effect and the slope effect, thus increasing this study's statistical power. As a fifth strength we would like to mention the contribution of our study to the existing literature, by presenting a way of incorporating whole bodily movements in whole-classroom lesson activities.

6.3. Conclusions and implications for education

The aim of this study was to incorporate (physical) experiences during graphing activities as embodied support in mathematics lessons in order to positively contribute to fifth-grade students' understanding of distance-time graphs. This study showed that the used activities resulted in higher levels of graphical reasoning, thus demonstrating the usefulness of incorporating graphing activities in the primary school mathematics classroom. Additionally, this study showed the added value of physical activities, as whole bodily movements in front of the motion sensor, on students' graphical reasoning. The current study adds to a growing body of evidence that physical experiences are indeed helpful for mathematics learning in general and graphical understanding in particular. Yet, what exactly caused this growth is something further research could explore.

Even though on the basis of this study we cannot make strong statements, we do think our study has some implications for graphing motion in primary school mathematics classrooms. First, through carefully designed lesson activities involving problem situations situated in a real-world context, capitalizing on students' intuitive understandings of representing motion, students' graphical reasoning can be improved. Second, our study shows that it is possible to implement embodied activities, that are activities enriched with immediate whole-bodily motion experiences, in an authentic classroom setting (see also Deniz & Dulger, 2012), which adds to research investigating practical applications of embodied cognition approaches for education and learning. In this respect, our study confirms findings from previous research into embodied mathematics learning showing the feasibility of incorporating these type of physical bodily-based activities in whole classrooms.

Acknowledgements

This research was supported by grant 405-14-303 from The Netherlands Initiative for Education Research (NRO) for the Beyond Flatland project.
We would like to thank the students, teachers and the schools for their participation in this study.

References

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics*, *14*(3), 24–35.

Arzarello, F., Pezzi, G., & Robutti, O. (2007). Modelling body motion: An approach to functions using measuring instruments. In W. Blum, P. L. Galbraith, H. W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education* (pp. 129–136). Boston, MA: Springer.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science, 2*, 716–724. doi:10.1111/j.1756-8765.2010.01115.x.

Berg, C. A., & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments. *Science Education*, *78*, 527–554. doi:10.1002/sce.3730780602.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 17–66). Dordrecht: Springer.

Boote, S. K. (2014). Assessing and understanding line graph interpretations using a scoring rubric of organized cited factors. *Journal of Science Teacher Education*, *25*, 333–354. doi:10.1007/s10972-012-9318-8.

Botzer, G., & Yerushalmy, M. (2008). Embodied semiotic activities and their role in the construction of mathematical meaning of motion graphs. *International Journal of Computers for Mathematical Learning*, *13*, 111–134. doi:10.1007/s10758-008-9133-7.

Brasell, H. M., & Rowe, M. B. (1993). Graphing skills among high school students. *School Science and Mathematics*, *93*, 63–70. Doi:10.1111/j.1949-8594.1993.tb12196.x.

Clement, J. (1985). Misconceptions in graphing. In L. Streefland (Ed.), *Proceedings of the Ninth International Conference for the Psychology of Mathematics Education* (pp. 369–375). Utrecht: Utrecht University.

Deniz, H., & Dulger, M. F. (2012). Supporting fourth graders' ability to interpret graphs through real-time graphing technology: A preliminary study. *Journal of Science Education and Technology*, *21*, 652–660. doi:10.1007/s10956-011-9354-8.

diSessa, A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10*(2), 117–160.

Doorman, L. M., & Gravemeijer, K. P. E. (2009). Emergent modeling: discrete graphs to support the understanding of change and velocity. *ZDM – The International Journal on Mathematics Education*, *41*, 199–211. doi:10.1007/s11858-008-0130-z.

Duijzer, C., Van den Heuvel-Panhuizen, M., Veldhuis, M., Doorman, M., & Leseman, P. (2019).

Embodied learning environments for graphing change: A systematic literature review. *Educational Psychology Review*. *31*(3), 597–629. doi:10.1007/s10648-019-09471-7

English, L. D. (2016). STEM education K-12: perspectives on integration. *International Journal of STEM Education*, *3*(3). 1–8. doi:10.1186/s40594-016-0036-1.

Ferrara, F. (2014). How multimodality works in mathematical activity: Young children graphing motion. *International Journal of Science and Mathematics Education*, *12*. 917–939. doi:10.1007/s10763-013-9438-4.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*, 124–158. doi:10.2307/749671.

Godwin, S., & Sutherland, R. (2004). Whole-class technology for learning mathematics: the case of functions and graphs. *Education, Communication & Information*, *4*(1), 131–152. doi:10.1080/1463631042000210953

Goldman, A. I. (2012). A moderate approach to embodied cognitive science. *Review of Philosophy and Psychology*, *3*(1), 71–88. doi:10.1007/s13164-012-0089-0

Heck, A., Kędzierska, E., & Ellermeijer, T. (2009). Design and implementation of an integrated working environment for doing mathematics and science. *Journal of Computers in Mathematics and Science Teaching*, *28*(2), 147–161.

Honey, M., Pearson, G., & Schweingruber, A. (2014*). STEM integration in K-12 education: Status, prospects, and an agenda for research*. Washington, DC: National Academies Press.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86–104. doi:10.1037/a0034008

Johnson, H.L., McClintock, E., & Gardner, A. (in press). Leveraging difference to promote students' conceptions of graphs as representing relationships between quantities. In: U.T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in MathematicsEducation*. Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS Rekenen-Wiskunde voor groep 3 tot en met 8*. Arnhem: Cito.

Kaput, J. J., & Roschelle, J. (2013). The mathematics of change and variation from a millennial perspective: New content, new context. In S.J. Hegedus, & J. Roschelle (Eds.), *The SimCalc Vision and Contributions* (pp. 13–26). Dordrecht: Springer.

Krabbendam, H. (1982). The non-qualitative way of describing relations and the role of graphs: Some experiments. In G. van Barnveld, & H. Krabbendam (Eds.), *Conference on Functions* (pp. 125–146). Enschede: SLO.

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

Lingefjärd, T., & Farahani, D. (2017). The elusive slope. International Journal of Science and Mathematics Education, 1–20. doi:10.1007/s10763-017-9811-9

Matuk, C., Zhang, J., Uk, I., & Linn, M. C. (2019). Qualitative graphing in an authentic inquiry context: How construction and critique help middle school students to reason about cancer. *Journal of Research in Science Teaching*. 56(7), 905–936. doi:10.1002/tea.21533

Mishra, P., & Henriksen, D. (2018). Creativity, technology & education: exploring their convergence. Cham: Springer.

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*, 369–383. doi: 10.1002/tea.3660240408.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics.* Reston, VA: NCTM.

Nemirovsky, R., Kelton, M. L., & Rhodehamel, B. (2013). Playing mathematical instruments: Emerging perceptuomotor integration with an interactive mathematics exhibit. *Journal for Research in Mathematics Education, 44*, 372 415.doi:10.5951/jresematheduc.44.2.0372.

Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction, 16*, 119–172. doi:10.1207/s1532690xci1602_1.

Pouw, W. T., Van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, *26*, 51–72. doi:10.1007/s10648-014-9255-5.

Radford, L. (2009b). “No! He starts walking backwards!”: interpreting motion graphs and the question of space, place and distance. *ZDM–The International Journal on Mathematics Education, 41*, 467–480. doi:10.1007/s11858-009-0173-9

Radford, L. (2009b). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics, 70*, 111–126. doi:10.1007/s10649-008-9127-3.

Radford, L. (2014). Towards an embodied, cultural, and material conception of mathematics cognition. *ZDM – The International Journal on Mathematics Education*, *46*, 349–361. doi:10.1007/s11858-014-0591-1.

Radford, L., Cerulli, M., Demers, S., & Guzman, J. (2004). The sensual and the conceptual: Artifact mediated kinesthetic actions and semiotic activity. In M. J. Hoines, & A. B. Fuglestad (Eds.), *Proceedings of the 28 Conference of the International Group for the Psychology of Mathematics Education* (pp. 73–80). Bergen: Bergen University College.

Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist*, *44*, 162–169. doi:10.1037/0003-066X.44.2.162.

Robutti, O. (2006). Motion, Technology, Gestures in Interpreting Graphs. *International Journal for Technology in Mathematics Education, 13*(3), 117–125.

Roschelle, J., Kaput, J., & Stroup, W. (2000). SimCalc: Accelerating student engagement with the mathematics of change. In M. J. Jacobsen, & R.B. Kozma (Eds.), *Learning the sciences of the 21st century: Research, design, and implementing advanced technology learning environments*. (pp. 1–43). Hillsdale, NJ: Earlbaum.

Roth, W. M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research, 71*, 365–392. doi:10.3102/00346543071003365.

Schoenfeld, A. H. (2016). *An introduction to the Teaching for Robust Understanding (TRU) Framework. Resource document.* Berkeley, CA: Graduate School of Education. http://map.mathshell.org/trumath/intro_to_tru_20161223.pdf. Accessed 5 April 2017.

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications*, *3*(6), 1–10. doi:10.1186/s41235-018-0092-9

Streefland, L. (1985). Wiskunde als activiteit en de realiteit als bron [Mathematics as an activity and the reality as a source]. *Tijdschrift voor Nederlands Wiskundeonderwijs (Nieuwe Wiskrant), 5*(1), 60–67.

Zaslavsky, O., Sela, H., & Leron, U. (2002). Being sloppy about slope: The effect of changing the scale. *Educational Studies in Mathematics*, *49*(1), 119–140. doi:10.1023/A:1016093305002