

The Link between Microdevelopment and Long-Term Learning Trajectories in Science Learning

Steffie van der Steen^a Henderien W. Steenbeek^b
Ruud J.R. Den Hartigh^b Paul L.C. van Geert^b

^aDepartment of Pedagogical and Educational Sciences, University of Groningen, Groningen, The Netherlands; ^bDepartment of Psychology, University of Groningen, Groningen, The Netherlands

Keywords

Dynamic systems · Microdevelopment · Skill theory · Variability · Understanding · Air pressure

Abstract

Long-term learning trajectories evolve through microdevelopmental sequences (i.e., short-term processes of change during learning tasks) and depend on variability during and across learning tasks. The aim of this study is to examine the coupling between short-term teacher-student dynamics and students' long-term learning trajectories, thereby providing empirical support for the link between the short- and long-term time scale in cognitive development. For 31 students (ages 3–5 years) from regular and special education, five teacher-student interactions during science tasks were filmed and coded in real time with regard to the student's level of understanding and the teacher's support throughout the task. A hierarchical cluster analysis resulted in four different learning trajectories over the course of 1.5 years, labeled as a high-scoring, mid-scoring, fluctuating, and low-scoring group of students. When connecting these trajectories to microdevelopmental data, the interactions of the high-scoring students were characterized by more moment-to-moment variations in the teacher's support and student's level of understanding, while the low-scoring group had the least variability compared to the other groups. Students with emotional and behavioral disabilities were represented across all learning trajectories, despite frequent accounts in the literature on their significant academic delays.

© 2019 The Author(s)
Published by S. Karger AG, Basel

KARGER

E-Mail karger@karger.com
www.karger.com/hde

© 2019 The Author(s)
Published by S. Karger AG, Basel

Karger
Open access

This article is licensed under the Creative Commons Attribution-Non-Commercial-NoDerivatives 4.0 International License (CC BY-NC-ND) (<http://www.karger.com/Services/OpenAccessLicense>). Usage and distribution for commercial purposes as well as any distribution of modified material requires written permission.

Steffie van der Steen
Department of Pedagogical and Educational Sciences
Faculty of Behavioural and Social Sciences, University of
Groningen, NL-9712 TS Groningen (The Netherlands)
E-Mail s.van.der.steen@rug.nl

Long-term development evolves through a series of microdevelopmental sequences and depends on variability during and across learning tasks (Granott, 2002; Van Geert, 1994). A better understanding of the short-term processes of change would therefore provide clues on how to optimize curricula, teaching methods, and learning materials (Granott & Parziale, 2002). Studies on such short-term processes are called microdevelopmental studies, which enable researchers to examine the process of change in students' understanding while they learn. Some of these studies take a dynamic systems approach and focus on real-time patterns of change during the learning process, often during a single task or a series of tasks (Fischer & Granott, 1995; Guevara, Van Dijk, & Van Geert, 2016; Montes, Van Dijk, Puche, & Van Geert, 2017; Van der Steen, Steenbeek, Van Dijk, & Van Geert, 2014; Van Vondel, Steenbeek, Van Dijk, & Van Geert, 2017). Researchers have claimed that the microdevelopmental approach can identify developmental patterns and detect transitions and changes in understanding (e.g., Brock & Taber, 2017; Granott & Parziale, 2002).

Few empirical studies, however, have specifically focused on the coupling of microdevelopmental data to long-term development over the course of several months or years (for an exception in science learning, see Guevara, Cox, Van Dijk, & Van Geert, 2017; Montes et al., 2017; Tytler, 1998; Van der Steen et al., 2014). Moreover, while the student and his or her (social) context are dynamically intertwined (Brown, Collins, & Duguid, 1989; Rogoff & Lave, 1984; Van der Steen, Steenbeek, & Van Geert, 2012a; Van Geert, 2002), few studies present data of teacher-student interactions linked to students' learning processes over a longer developmental period. Such information is, however, important to understand how differences in learning trajectories emerge. For example, several studies have shown that students with special needs, and especially those who have emotional or behavioral problems, generally perform worse at school (Lane, Barton-Arwood, Nelson, & Wehby, 2008; Reid, Gonzalez, Nordness, Trout, & Epstein, 2004), which is often related to a negative dynamics between teacher and student (Steenbeek, Jansen, & van Geert, 2012; Steenbeek & Van Geert, 2013; Wehby, Lane, & Falk, 2003). Especially with regard to these students, understanding how real-time teacher-student interactions contribute to their learning trajectory over time matters (Steenbeek et al., 2012).

The current study focuses on microdevelopmental data of students' understanding and teacher's instructional support during five hands-on science tasks about air pressure and connects these data to students' learning trajectories over the course of 18 months. Creating such a coupling between short-term teacher-student dynamics and long-term learning trajectories is crucial for a thorough understanding of how knowledge is transferred over time, and also provides important clues on how to optimize learning trajectories (Granott & Parziale, 2002; Van Geert & Steenbeek, 2005b). Hence, while this paper covers science learning, the study has important value for all studies concerned with the development of students' understanding over time, regardless of its specific nature or domain. We will proceed by discussing the theory of dynamic systems, the theoretical base that warrants the connection between micro- and long-term development. We then focus on variability, an important mechanism by which dynamic systems change over time so that new patterns can emerge. Lastly, teacher-student interactions are discussed from a dynamic point of view, and we highlight students for whom these interactions may be more difficult: students with emotional and behavioral disabilities.

The Theory of Dynamic Systems

The theory of dynamic systems is an approach to the description and explanation of change (Van Geert & Fischer, 2009). It provides a powerful theoretical base and can serve as a metatheory to connect micro- and macrodevelopmental processes (e.g., Granic & Patterson, 2006; Lichtwarck-Aschoff, Van Geert, Bosma, & Kunnen, 2008; Newell, Liu, & Mayer-Kress, 2001; Thelen & Corbetta, 2002; Thelen & Smith, 1994; Van Geert, 1994, 1998a, b; Van Geert & Steenbeek, 2005a). To be more specific, both micro- and macrodevelopment share similar dynamic systems principles, such as self-organization, emergence (the process by which new skills or insights are created), nonlinear growth, stage-like transitions over time, variation within a certain developmental range, and multimodality, that is, the simultaneous existence of several developmental levels (e.g., Thelen & Smith, 1994; Van Geert, 1994, 1998a).

Development starts with the behavior of a system (e.g., the student), which is the product of dynamically interacting components, including those that are part of the system (aptitude, skill level) and those that are part of the environment (teacher support, learning task properties). These components are constantly changing over time and self-organize, meaning that there is an unorchestrated continuous interaction (Den Hartigh, Van Dijk, Steenbeek, & Van Geert, 2016; Thelen & Corbetta, 2002; Van Geert, 1991, 1998a). Development can therefore be seen as a process of intertwining person-context dynamics (Thelen & Smith, 1994), meaning that the social (e.g., the teacher) and material context (e.g., learning task) cannot be viewed separately, as external influences (Lee & Karmiloff-Smith, 2002; Van der Steen et al., 2012a). At any moment in time, one component (e.g., the student) affects the other (e.g., the teacher) and vice versa. This results in a continuous person-environment loop, which Fogel and Garvey (2007) called a “cooperative unit.” Hence, a student’s understanding of scientific phenomena is an active process of what the student constructs in interaction with (not just within) a specific environment (Van Geert & Fischer, 2009). This resembles the ecological view on cognition, according to which the organism-environment system should be the unit of analysis (e.g., Turvey, 2009; Turvey & Shaw, 1999).

Human behavior exists in and changes over time. An individual’s behavior in the future, such as a more elaborate understanding of a science task, develops out of the current understanding of the student (Den Hartigh et al., 2016; Van Geert, 1991). This iterative process in which “after is explained by before” (Van Geert & Steenbeek, 2005a, p. 408) continues throughout life. Since development is the product of multiple interacting components that self-organize and produce patterns of change by means of an iterative process, it is considered to be nonlinear (Thelen & Corbetta, 2002; Van Geert, 1998b). That is, there is no fixed, one-to-one relationship between the magnitude of a cause and the magnitude of an effect (e.g., more instructional support causes higher levels of understanding). In other words, a nonlinear process entails that similar influences can shape development in a number of ways. Some systems are sensitive to relatively small changes in its components, leading to a reconfiguration such as a transition to a more advanced understanding (Thelen & Corbetta, 2002). Hence, an important consequence of nonlinearity is that differences between and within students are partly caused by factors that are intrinsic to the developmental process (Van Geert, 1998b).

Interestingly, there is a circular relationship between microdevelopment (e.g., changes in understanding during learning tasks) and macrodevelopment (e.g., changes in understanding over the course of months or years). While real-time processes

produce patterns of change that shape our development on the long term, established patterns of development also influence real-time iterative processes (Lewis, 2002). For example, if a child discovers how air flows through a syringe, this contributes to his/her repertoire of understanding air pressure (micro to macro), but also influences how he/she encounters a new air pressure task at a later moment in time (macro to micro). On both time scales, mechanisms of learning and development can be better understood when examining system variability over time (e.g., Lichtwarck-Aschoff, Kunnen, & van Geert, 2009; Steenbeek et al., 2012; Thelen, Ulrich, & Wolff, 1991; Van Geert & Van Dijk, 2002).

System Variability. Change in the understanding of a student can only happen if the configuration of the system's components is flexible. This means that components should be able to reassemble in order to break up maintaining patterns so that new patterns can emerge (Siegler, 2007; Thelen & Corbetta, 2002). Hence, variability enables the system to grow and change by providing flexibility (Hayes & Strauss, 1998). Understanding the mechanisms of learning and development therefore requires the study of variability and assessing its patterns (Yan & Fischer, 2002).

Especially on the short-term time scale, developmental patterns are characterized by high variability and smooth growth is rare (Yan & Fischer, 2002). Typically, students vary within a so-called developmental range, the boundaries of which can be described as upper and lower attractors. Fischer and colleagues (Fischer, 1980; Fischer & Bidell, 2006; Fischer & Granott, 1995; Granott, 2002; Yan & Fischer, 2002) defined the lower attractor as the current level the student can reach without support and the upper attractor as the level the student can reach with high-quality support. In this way, the range reflects a Vygotskian (1978) zone of proximal development (Van Geert, 1998a; Van Geert & Steenbeek, 2005b). Variations within the developmental range are therefore not only due to changes in the student's understanding and skills, but also to changing task and contextual constraints (Fischer & Bidell, 2006; Yan & Fischer, 2002).

On a macrodevelopmental scale, Van Geert (1998a) speaks about synchronous multimodality in this regard, i.e., the simultaneous existence of different levels of understanding. The Piagetian notion of "horizontal décalage," a term coined to describe the situation in which children function on two or more levels at the same time (Piaget, 1947/2001; Van Geert, 1998a), is related to this. This multimodality can be an indication of a developmental transition to a more advanced level, and during the transition phase the system regularly switches between the new and old levels (Howe & Lewis, 2005; Van der Steen et al., 2012a; Van Geert, 1994). Thus, the occurrence of varying levels of students' understanding at roughly the same time has its roots in the principles of dynamic systems and can be indicative for a transition to a higher level of understanding.

Several studies have shown that fluctuations in understanding even occur during tasks, on the microdevelopmental time scale (Fischer & Granott, 1995; Granott, 2002; Siegler, 1995; Yan & Fischer, 2002). Siegler's microdevelopmental study (1995), for instance, has shown that students can have multiple ways of thinking at the same time, and that the composition and distribution of their different thinking strategies changes over time. Yan and Fischer (2002) suggest that the magnitude of variability in learning not only depends on developmental transitions, as discussed above, but also on the student's relative position in the learning process. At the start of the learning process, students often show more extreme fluctuations, as they build their skills, but also have

trouble to maintain the higher levels they achieve. When students almost master a certain task or thinking level, less extreme and slower fluctuations occur (Van Geert & Van Dijk, 2002). Since variability in understanding occurs both at the micro- and macrodevelopmental time scale and reflects an important part of the developmental process, we need a measure of understanding that is sensitive enough to capture these fluctuations (Brock & Taber, 2017; Thelen & Corbetta, 2002) on both the micro- and macrodevelopmental time scale. Such a measure is grounded in skill theory.

Skill Theory

Skill theory is a cognitive developmental theory that focuses on the complexity and variations of children's thinking abilities (skills), and the way these are constructed (Fischer, 1980; Fischer & Bidell, 2006). The basis of the theory is a hierarchical scale to measure the complexity of skills, which can be used to assess understanding levels over a longer period of time, but also while students work on single tasks. The scale consists of several levels, grouped into three tiers of increasing complexity. The first tier consists of sensorimotor skills, in which perceptions are connected to actions or utterances. The second tier consists of representational skills, these are understandings that go beyond current observable perception-action couplings but are still based on them. The third tier consists of abstractions, which are general nonconcrete rules that also apply to other situations (Schwartz & Fischer, 2004). Within each tier, three levels of increasing complexity can be distinguished. The first one can be characterized as a single set, that is, a single action, a single representation, or a single abstraction. The second level is a relation between two of these sets, which is referred to as a mapping. The third level is a system of sets, which is a relation between two mappings, in which each mapping consists of a relation between single sets (Fischer & Bidell, 2006, see also the Method section).

Skill theory does not only describe and explain the development of skills on the long term, but can also be used to assess the microdevelopment of problem solving (Fischer, 1980; Fischer & Bidell, 2006; Fischer & Van Geert, 2014; Granott & Parziale, 2002; Schwartz & Fischer, 2004; Van Geert & Fischer, 2009; Yan & Fischer, 2002). For instance, when facing a new task, even skilled adults may show skill levels that are mostly sensorimotor at the beginning, building up to more elaborate levels. During tasks, people do not necessarily go through the skill cycles in a linear fashion. Instead, they may repeatedly build up skill levels and regress during the task, before they obtain their highest possible level (Yan & Fischer, 2002; see also the work of Clement, 2009, on imagistic simulation of experiments by scientists). Since skill theory focuses on the hierarchical complexity of skills rather than their content, the scale enables researchers to compare skills, such as students' understanding of scientific phenomena, across multiple time points, contexts, persons, and age ranges (Parziale & Fischer, 1998). Skill theory is therefore especially suitable to compare individual pathways within and across tasks (Fischer, Rose, & Rose, 2007).

Learning and Development

So far, our discussion of microdevelopment and long-term learning trajectories from a dynamic systems point of view has been intertwined. An important reason for

this is that studies of microdevelopment show that the dynamic characteristics of development are very similar to those of learning processes, and vice versa (Granott & Parziale, 2002). From a dynamic systems point of view, learning and development are therefore neither similar nor dissimilar, but two sides of the same continuum (Flynn & Siegler, 2007; Van Geert, 1995). As a result, short-term changes (microdevelopment) can be analyzed within the same framework as long-term changes (Fischer & Granott, 1995). Similar to learning, developmental processes are characterized by variability, and regressions to earlier levels occur (Granott, Fischer & Parziale, 2002). Learning processes, on the other hand, show characteristics of development, such as trends of growth (Gelman, Romo, & Francis, 2002), restructuring of knowledge, and shifts to higher levels of thinking (Granott, 1998; Granott & Parziale, 2002; Van Geert, 1998a). Learning can be achieved without external support (Granott & Parziale, 2002) but is often associated with structured interactions with more competent others, such as a teacher. In the next section, we further explore teacher-student interactions from a dynamic systems perspective.

Teacher-Student Interactions

Recall that from a dynamic systems perspective, the behavior of the student emerges through the processes of interaction between student, teacher, and material problem context. (e.g., Thelen & Smith, 1994; Steenbeek et al., 2012). This means that students' understanding emerges out of a series of actions they perform within a context and guided by their educator (Van Geert, 1998a; Van Geert & Steenbeek, 2005a; Steenbeek, van Geert, & van der Steen, 2017). Teacher-student interactions organize toward certain patterns of understanding during tasks (microdevelopment), which eventually evolve toward stable attractors on a longer time scale (Halley & Winkler, 2008).

Teacher-student interactions are generally studied at the classroom level or dyadic level (Vandenbroucke, Spilt, Verschueren, Piccinin, & Baeyens, 2018). Although factors such as classroom management, the social climate, and lesson structure are important elements that contribute to students' understanding, we focus on the dyadic level and more specifically on the instructional support that teachers provide their students. From a dynamic systems perspective, instructional support is also an example of a dynamic phenomenon, which self-organizes and emerges as a consequence of the interactions taking place between students and educators in particular problem contexts, given the histories of students and teacher in terms of their skills, expectations, habits, and knowledge (Van der Steen et al., 2012a). Instructional support often differs for each student and influences students' learning outcomes to a great extent (Hamre & Pianta, 2005). This is especially the case when support is intentional, focused, and contains feedback loops (Hamre & Pianta, 2005; Torgesen, 2002), for example when teachers use strategies like questioning, giving feedback, and explaining. Several authors indicate that high-quality instructional support should encourage students to take an active role in their learning process and should not be too directive (Ryan & Deci, 2000; Van Vondel et al., 2017). Especially when teachers use open-ended questions, students are given the opportunity to actively explore and construct their understanding (Topping & Trickey, 2014; Van Vondel et al., 2017). That said, instructional support may work out differently for different students, for example those with special needs (Wehby et al., 2003).

Several studies have shown that students with special needs, and especially those who have emotional or behavioral problems, generally perform worse at school (Lane et al., 2008; Reid et al., 2004). While the specific difficulties that these students encounter may differ depending on their diagnosis, they do resemble each other in that they generally display significant academic delays across all subject areas (Reid et al., 2004; Van der Steen, Steenbeek, Wielinski, & Van Geert, 2012b), which do not improve over time. Teachers often feel challenged when educating these children, which may result in a negative dynamics of teacher-student interactions (Steenbeek et al., 2012; Wehby et al., 2003). Wehby et al. (2003) indicate that these negative dynamics can either be caused by the students' behavioral problems interfering with the teaching, or by an unintentional difference in teacher behavior when confronted with students with emotional and behavioral disabilities (e.g., less praise for appropriate behavior) (Dunlap et al., 2006). However, research also shows that students with emotional and behavioral disabilities are able to perform at similar levels as their typically developing peers on educational science tasks when adequate instructional support is offered that is tailored to the student's needs (Van der Steen et al., 2012b).

Current Study

The aim of this study is to examine the coupling between short-term teacher-student dynamics and students' long-term learning trajectories, thereby providing empirical support for the link between the short- and long-term time scales in cognitive development. We investigated how 31 young students develop their understanding of air pressure during science tasks while they are guided by a teacher. The developmental trajectories are represented in the form of short- and long-term time series on a skill-theoretical scale. Students were 3–5 years old at the start of the study, and about half of them were recruited at special educational settings (designated schools for students with emotional and behavioral disabilities). Teacher-student interactions are central to this study, as we take both quantitative microdevelopmental data of the student's understanding (measured using a scale based on skill theory), as well as quantitative microdevelopmental data of the teacher's instructional support while they interact during five air pressure tasks, administered over the course of 18 months. In this study, we take a dynamic systems approach to microdevelopment (Lee & Karmiloff-Smith, 2002). That is, we view students' understanding as the product of interrelated and mutually interacting components, which reside both within the student as well as in the environment (Den Hartigh et al., 2016; Steenbeek et al., 2012; Thelen & Smith, 1994; Van Geert, 1998a). This means that we expect higher levels of understanding to emerge as a result of changes in the interaction between these components.

We start by looking at students' long-term development, that is, the global development of their understanding, and investigate whether different learning trajectories emerge. Our first research question is therefore: How do young students develop their understanding of science tasks over 18 months, and can we distinguish different learning trajectories, that is, different groups of learners? We then connect these long-term developmental patterns to the microdevelopment of these students during the five air pressure tasks. That is, can we relate the different learning trajectories to microdevelopmental interaction patterns during the science tasks? Specifi-

cally, we start by examining whether teacher-student interactions differ for the different groups of learners with regard to concrete student behaviors and teacher support (research question 2a). As indicated above, researchers have argued that long-term development depends on variability during and across learning tasks (Granott, 2002; Van Geert, 1997). Variations in understanding are, however, embedded within the social environment (Brown et al., 1989; Lave & Wenger, 1991; Van der Steen et al., 2012a). We therefore proceed by examining whether the teacher-student interactions of the groups of learners differ with regard to the moment-to-moment variability of the teacher's instructional support and the student's understanding levels (research question 2b). In this way, the study connects students' long-term learning trajectories to their microdevelopment, thereby combining multiple timescales in one study (Brock & Taber, 2017), in order to provide an important link that shows how understanding is formed (Granott & Parziale, 2002).

Method

Participants

The participants were 31 Dutch primary school students, of which 17 (10 male students, 7 female students) were typically developing students, and 14 (12 male students, 2 female students) were students enrolled in special educational settings. The lower percentage of female students in the special education group is in concordance with an overall lower percentage of female students in special schools (Reid et al., 2004). We recruited 3- to 5-year-old students, who were enrolled in (special) daycare centers or kindergarten at regular or special primary schools. Within these schools and centers, parents were asked if their children could participate in a longitudinal study on their understanding of science tasks. All students whose parents provided a written consent were included in the study. The study was approved by the ethics committee of the host institution.

The mean age of the students enrolled in regular education was 4.13 years ($SD = 0.96$); the students in special education were slightly older, with a mean age of 4.75 ($SD = 1.02$). The special schools' student population in the Netherlands is quite diverse, that is, both children officially diagnosed with internalizing (e.g., autism spectrum disorders, anxiety disorders) as well as externalizing challenges (e.g., attention deficit hyperactivity disorder, oppositional defiant disorder) are enrolled in special schools and daycare centers, and are taught in the same classrooms. In our study, 64% of the students in these schools had externalizing challenges and 36% internalizing challenges. The students in the regular schools were not diagnosed with emotional or behavioral problems.

Materials

During each visit (five in total, over the course of 1.5 years) the children worked on a scientific task about air pressure. Other authors have argued that studying understanding from a microdevelopmental perspective requires different subsequent educational tasks with similar underlying concepts (Brock & Taber, 2017). Each new task therefore required extra manipulations or more elaborate thinking to explain the mechanism and scientific concepts of the task (Fig. 1). The task sequence started with a toy frog that could jump by means of squeezing a balloon attached to its inflatable legs. The task of the second visit involved the connection of two syringes of the same volume by a tube, whose pistons moved in opposite directions when manipulating. During the third visit syringes with different volumes were connected to explore differences in their functioning. During the fourth visit several syringes were used to operate a miniature version of an elevator, and during the fifth visit the carrying capacity of such a syringe elevator was explored using air and water as content of the syringes.

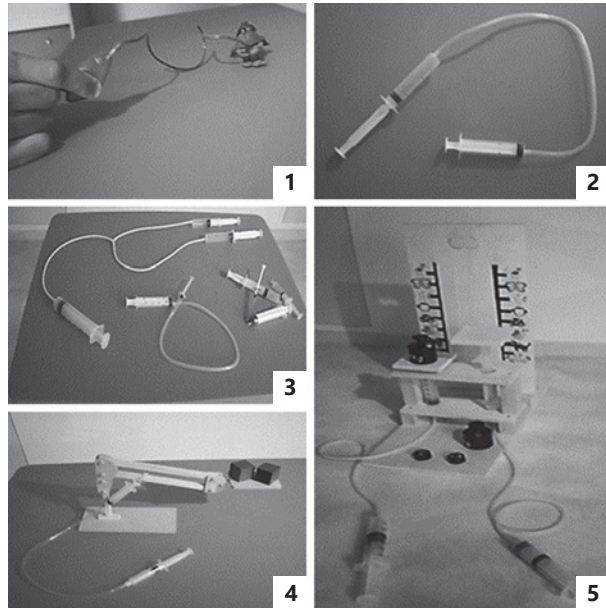


Fig. 1. The five air pressure tasks of this study.

Procedure

Visits. Every 3–4 months, students worked individually on the science tasks, guided by the first author of this paper (hereafter referred to as teacher), who designed the tasks and the task protocol. The protocol that was constructed for each task had the form of a semistructured interview, which guaranteed that all students were asked the same basic questions reflecting the core building blocks of the task and the incorporated scientific concepts. At the same time, the protocol left enough space for students to take initiative and show their understanding spontaneously, and for the teacher to provide instructional support when needed (see Fig. 2 for an excerpt of a protocol). For each task, the teacher showed the student the material and asked about its purpose and functioning. Afterward – regardless of whether the student answered the previous questions right – he or she was encouraged to explore the material. During this exploration, the teacher asked questions about the task’s functioning and underlying mechanisms, such as “Why does the piston of the other syringe go up when you push the piston of this syringe?” Besides the questions listed in the protocol, the teacher’s instructional support consisted of feedback, i.e., asking follow-up questions and providing encouragement to help students to elaborate on their answers, explanations, and directing the student’s focus when parts of the tasks were unclear to the student. The teacher used instructional support at her own discretion to help students to verbalize their highest level of understanding of the task at that certain moment. No other learning objectives were formulated. Follow-up questions and encouragement were used when the child’s answer was not immediately clear or complete. Even though students’ answers were challenged sometimes by means of these follow-up questions, the teacher never indicated whether answers were right or wrong. When the student could not give an answer, the teacher proceeded with the next question of the protocol. Each task took approximately 15–20 min. All interactions took place within students’ schools or daycare centers and were recorded on video. The teacher had no information about the students’ academic performance or diagnosis (if applicable) when she administered the tasks.

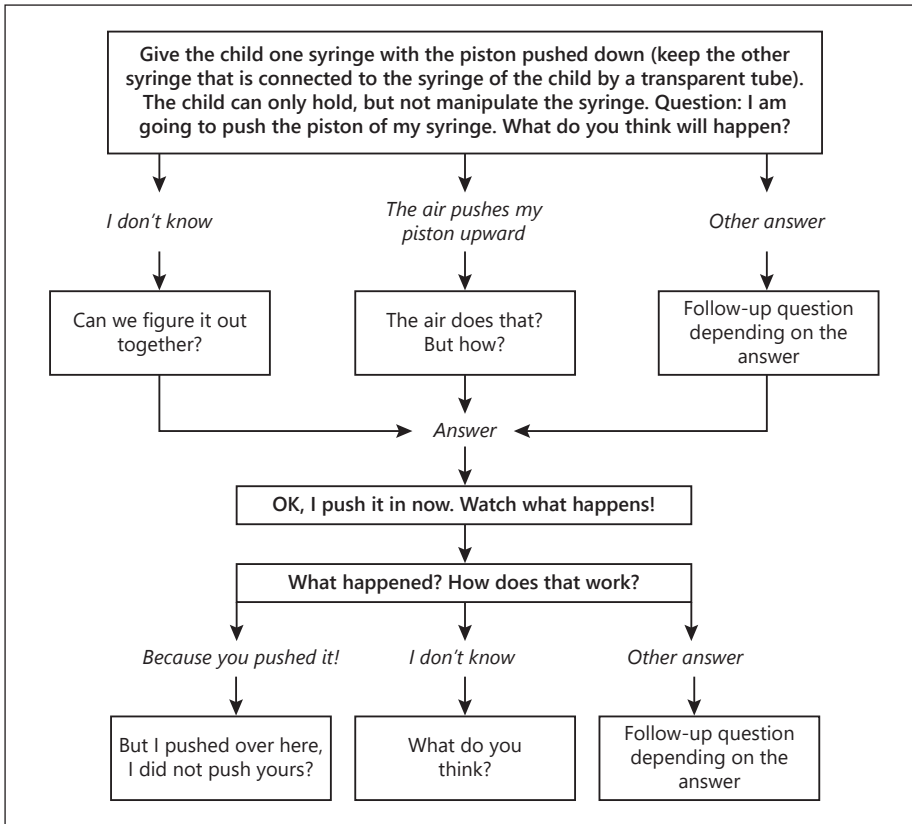


Fig. 2. First part of the protocol of the second visit.

Coding of Observational Data. Using the computer program MediaCoder (Bos & Steenbeek, 2006), the recordings of the five visits were coded per utterance and in real time using a standardized codebook (for previous work with this codebook, see Den Hartigh et al., 2014; Den Hartigh, Van der Steen, Hakvoort, Frencken, & Lemmink, 2018; Van der Steen et al., 2012a, b, 2014). The first step in the coding procedure was the determination of the moment when utterances of the student and the teacher started and ended. Subsequently, the teacher's utterances were coded as: descriptive, predictive, and explanatory questions; encouragement and follow-up questions; explanations and directing the student's focus; procedural remarks, and off-task utterances. Descriptive questions focused on the current observations ("What do you think this is?"), predictive questions focused on future states ("What do you think will happen if we push the piston of this syringe?"), and explanatory questions focused on the task's mechanism ("How do you think this works?"). Encouragement was coded when short encouraging words were used by the teacher ("ok," "go on"). Follow-up questions were coded when the teacher asked for additional information in relation to the child's previous answer. For example, when the child said "The air flows through the tube," the teacher could ask "And then what happens?" Explanations consisted of longer utterances of the teacher in which previous words or observations of the student were more explicitly clarified or summarized, for example: "So we saw that when we push in this piston, the other one goes backward." Procedural remarks were given to explicate the task administration, without focusing on the task content, such as: "This is the first task we are going to look at." Other comments that were not related to the task or the task administration were considered off-task.

Table 1. Descriptions and examples of the understanding levels, based on skill theory

Level	Description of answer (example)
1 Single sensorimotor set	Single characteristic or perception of the task (“The piston goes up”)
2 Sensorimotor mapping	Link between single task characteristics; simple comparisons (“We can push the piston into this syringe”)
3 Sensorimotor system	Observable causal relations (“If I push the piston of this syringe, then the piston of the other syringe moves”)
4 Single representation	Coupling two causal relations; not directly observable relations; simple predictions (“Air causes the piston of the syringe to move”)
5 Representational mapping	Predictions and explanations in terms of a relation between two single representations (“The piston pushes the air, and because of the tube connecting the two syringes, the other piston gets pushed out by the air”)
6 Representational system	A coupling between two representational mappings (“This syringe contains air, and if I push its piston, the air is pushed through the tube, and pushes that piston upward. When I push the other piston, the same mechanism causes the first one to go up”)
7 Single abstraction	A general concept that goes beyond (representations of) the current task (“When I push this piston, the air is compressed....”)

Students’ utterances were classified into: descriptive (observations), predictive (hypotheses), and explanatory answers; requests to manipulate the task, questions about the task, and off-task utterances. After this initial classification, we combined students’ coherent descriptive, predictive, and explanatory answers into meaningful units to enable further coding of the understanding levels. That is, we combined utterances about the same topic that were verbalized successively and without any interruption. The unit ended when the next utterance of the student fell into a different category, or when the researcher asked another question. However, if the researcher only encouraged the student to tell more about the same topic (“Ok, go on...”), the unit would continue.

The complexity levels of students’ answers within a unit were then determined using a scale based on skill theory (Fischer, 1980; Fischer & Bidell, 2006). This scale enables researchers to filter out the complexity of the task-related utterances in this study, making it possible to investigate the development of understanding during a task (see also Van der Steen et al., 2012a, b, 2014). The levels assigned in our study form an ordinal scale, ranging from single sensorimotor sets (level 1) to single abstractions (level 7); see Table 1 for level descriptions and examples. We did not assign complexity levels to incorrect task-related utterances.

Interrater Reliability. To ensure interrater reliability, a standardized coding protocol was used. For each step of coding (teacher and student utterance categories, units of students’ successive task-related utterances, understanding levels), 10 raters went through a training by coding three videos of 15 min. The codes of the third video were then compared to the codes given by the first author, who constructed the coding protocol. When calculating percentages of agreement, both the given codes as well as the timing were taken into account. Differences in timing

of less than 2 s were considered acceptable. The average percentages of agreement were 83% for the teacher and student categories (range 80–93%; $p < 0.01$), 87% for constructing units of successive student answers (range 80–100%; $p < 0.01$), and 84% for assigning understanding levels (range 78–92%; $p < 0.01$). The p values were calculated using Monte Carlo permutation tests (see below for an explanation).

Analysis

Differences in the Long-Term Development of Understanding (RQ 1). In order to provide a picture of the long-term development of students' understanding of air pressure over the course of 18 months, we calculated the proportion of representational and abstract levels of understanding for each visit (level 4 and higher) and for each student. The choice for this measure was based on the nature of the representational and abstract levels. Since these levels go beyond observable perception-action couplings (i.e., what students see), they reflect what children understand about the core topic underlying the tasks. We calculated the proportion of representational and abstract levels for each visit, resulting in five consecutive proportions for all 31 students, reflecting the longitudinal development of their understanding of these tasks over the course of 18 months.

To see whether our data set existed of groups of students with similar long-term trajectories, we performed a hierarchical agglomerative clustering (HAC) analysis in Tanagra 1.4.18 (Rakotomalala, 2005). We used the five proportions of representational and abstract levels for each student as input. HAC successively merges similar individuals into smaller groups, keeping the within-cluster variance at a minimum (Borgen & Barnett, 1987). The clustering starts by joining the two most similar subjects together, after which the next two most similar are merged, until coherent groups of students with similar developmental trajectories are formed. We explored the development of the resulting groups over time and calculated between-group differences with regard to the proportion of representational and abstract understanding levels at each visit.

To determine the significance levels of these differences, we used Monte Carlo permutation tests (Todman & Dugard, 2001). This statistical procedure can easily be applied to small samples and skewed distributions, while still producing reliable statistical results. Taking the sample distribution into account, a Monte Carlo test measures the probability that a difference is caused by chance. This is done by drawing many random samples (in this case 1,000) from the original data, after which one can determine how often the observed or a bigger difference occurs in these random samples. This number is divided by the number of samples (1,000), which produces a p value. After exploring the between-group differences in understanding levels, we also used Monte Carlo permutation tests to explore differences in group composition with regard to the distributions of demographic factors age, gender, and school type (regular or special schools).

Differences in Teacher Support and Student Behaviors (RQ 2a). We proceeded by investigating differences between the clusters of students with regard to the students' and teacher's behaviors during the tasks. For each teacher-student interaction we calculated the number of teacher utterances, and the proportions of the following measures indicating instructional support: (a) questions, (b) follow-up questions and encouragement, and (c) explanations and directing the student's focus. For each group of students (i.e., cluster) we then calculated the average number of these measures for each visit. Using Monte Carlo permutation tests, we examined differences in instructional support between the groups at each visit and for all visits combined. We then calculated the average number of student utterances, and the average proportions of (a) task-related answers and (b) student-initiated questions (requests to manipulate the task and asking questions about the task) for each visit and for each group. Using Monte Carlo permutation tests, we examined differences in student behavior between the groups at each visit and for all visits combined.

Variability in Real-Time Interaction Patterns (RQ 2b). Lastly, we examined differences between the groups in terms of the real-time interaction patterns between teacher and student. To investigate this relationship, we first recoded the teacher's behaviors to an ordinal scale reflecting the directive nature of the teacher's instructional support. Descriptive, predictive, and explana-

tory questions were coded as 3, a higher-level form of support, which was the least directive level that left the most room for the student to contribute to the interaction. Feedback in the form of encouragement and follow-up questions was coded as 2, being slightly more directive, as these were focused on a more delimited aspect of the task. Explanations and directing the student's focus were coded as 1, as these indicate that the teacher takes over the task, limiting the student's contributions the most.

We created time series of the (recoded) teacher's instructional support and the student's (skill theory) understanding levels with a sampling rate of 2 Hz for each student and for each visit. We smoothed these time series using a local regression technique (Loess; Jacoby, 2000). That is, we used a weighted least squares formula to fit each local polynomial in the graph based on the data of 45 adjacent seconds.

We then focused on the moment-to-moment relationship between fluctuations of the teacher's instructional support and the student's understanding levels by calculating the concordance coefficient at each data point, using this formula:

$$\sqrt{|a|(\Delta \text{ understanding} \times \Delta \text{ support})} \times \text{sign}(\Delta \text{ understanding} \times \Delta \text{ support}).$$

A positive value of this measure means that both variables show a similar direction of change between two consecutive time points. This could either mean that the teacher's support changes to higher quality (i.e., less directive, more room for the student's contributions) while the student's understanding levels rise, or that the teacher's support changes to a lower quality (more directive, less room for the student's contributions) while the student's understanding levels drop. A negative value of the concordance coefficient means that the change in the teacher's support and student's understanding levels is dissimilar, that is, either the support changes to a higher quality while the understanding level decreases, or the support changes to a lower quality while the understanding level increases. High absolute concordance values indicate bigger changes while values close to zero indicate a smaller change.

To further investigate the variability of instructional support and understanding within the real-time interaction patterns, we calculated the standard deviation of the concordance values for each student and each visit, and took the average for each group of students. Since the positive and negative concordance values indicate a different relationship between teacher support and students' understanding, we calculated the standard deviations separately for the positive and negative concordance values. We graphically displayed the variation in concordance by group and used Monte Carlo permutation tests to examine whether the positive and negative variations in concordance differed significantly between the groups, for all visits.

Results

Differences in the Long-Term Development of Understanding (RQ 1)

The HAC analysis resulted in four different groups of students (Fig. 3). In concordance with the observed patterns, we labeled these groups as high-scoring ($n = 3$), mid-scoring ($n = 12$), fluctuating ($n = 5$), and low-scoring ($n = 11$). In general, over 50% of the answers of the highest scoring group of students could be classified as single representation or higher, exceeding the other three groups. Their understanding levels peaked at visit 2, while there was a drop at visit 3. The mid-scoring group of students showed the same pattern over the five visits, but the proportions of representational and abstract understanding levels were lower compared to the highest scoring group. The fluctuating group showed a deviating pattern. Their proportions of representational and abstract understanding levels varied from 0 (visit 1) to 0.54 (visit 2). This means that over the course of 3 months, these children proceeded from

Table 2. Composition of the groups (age, sex, school type)

Group	Mean age (SD), months	Proportion of males	Proportion of special education
High ($n = 3$)	65.67 (7.23)	1.00	0.33
Mid ($n = 12$)	55.92 (11.86)	0.67	0.42
Fluctuating ($n = 5$)	47.40 (7.80)	0.40	0.20
Low ($n = 11$)	48.73 (11.16)	0.82	0.64

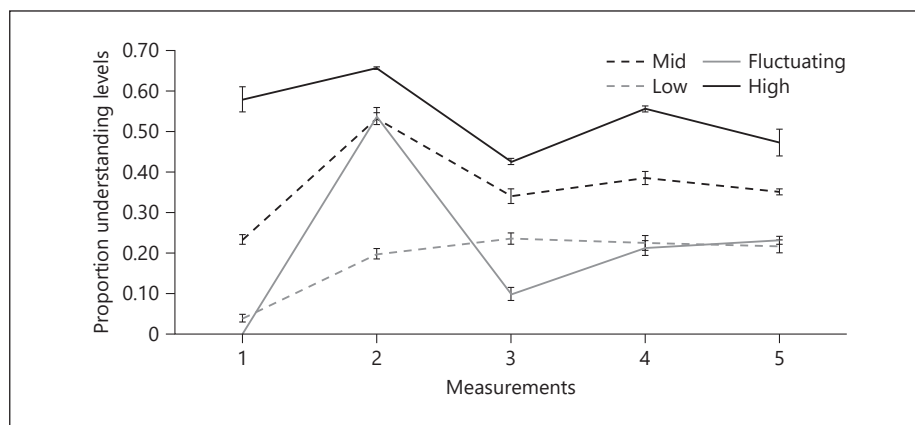


Fig. 3. Average pattern of representational and abstract understanding levels over the five visits, displayed by group: high-scoring ($n = 3$), mid-scoring ($n = 12$), fluctuating ($n = 5$), and low-scoring ($n = 11$). Error bars reflect the standard error.

only showing sensorimotor understanding levels to showing high understanding levels in more than 50% of their verbalizations. This proportion dropped to 0.1 at the third visit, and then rose again to roughly 0.20 during the final two visits. Lastly, the group with the lowest scoring students showed an increase of 0.16 from the first to the second visit, but then the representational and abstract understanding levels remained stable during the last three visits (Fig. 3).

Monte Carlo permutation tests revealed that the scores of the high-scoring group significantly differed from the low-scoring group on all visits ($p < 0.05$). There were also significant differences between the high-scoring and the fluctuating group ($p < 0.01$) on all visits, apart from the second visit ($p = 0.24$). In addition, the highest-scoring group significantly differed from the mid-scoring group at visit 1 ($p < 0.01$) and visit 4 ($p < 0.05$), but not at visit 2 ($p = 0.19$), visit 3 ($p = 0.17$) and visit 5 ($p = 0.08$). The mid-scoring group differed significantly from the low-scoring group on all visits ($p < 0.05$) and differed from the fluctuating group on all visits ($p < 0.05$), apart from the second visit ($p = 0.53$). The fluctuating and the low-scoring group differed significantly on the second visit ($p < 0.01$), at which the fluctuating group scored higher, and on the third visit ($p < 0.05$), at which the low-scoring group scored higher.

Table 3. Teacher utterances and proportion of support across all visits, by group

	Mean number of utterances	Mean proportion of questioning	Mean proportion of follow-up and encouragement	Mean proportion of explaining
High	149.47	0.21	0.37	0.09
Mid	181.52	0.23	0.32	0.10
Fluctuating	180.44	0.23	0.32	0.10
Low	178.89	0.25	0.28	0.10

Group Composition. Differences in group composition with regard to age, gender, and school type are listed in Table 2. On average, the high-scoring group was significantly older than the fluctuating ($p < 0.05$) and low-scoring ($p < 0.05$) groups. The age difference between the high-scoring and the mid-scoring group was not significant, as were the age differences between the other groups. The proportion of male students in the high-scoring group differed significantly from the fluctuating group ($p < 0.05$), but not from the other two groups. There was no significant difference in the proportion of male students between the other groups. The number of students with emotional and behavioral disabilities did not differ significantly between the groups, apart from the difference between the fluctuating (0.2) and the low-scoring (0.64) groups ($p < 0.01$).

Differences in Terms of Teacher Support and Student Behaviors (RQ 2a)

Teacher Instructional Support. Overall the teacher used significantly fewer utterances in interactions with children in the high-scoring group compared to the other groups (mean difference, $M_{\text{diff}} = 30.81$; $p < 0.01$) (Table 3). This was also true for the last three visits ($p < 0.05$). The other groups did not differ in terms of teacher utterances, apart from a single difference between the fluctuating and the low-scoring group on visit 1 ($M_{\text{diff}} = 22.27$, $p < 0.05$).

Interactions with students in the low-scoring group contained an overall higher proportion of questioning compared to the other groups ($M_{\text{diff}} = 0.03$; $p < 0.05$). This was, however, mostly caused by differences between the low-scoring group and the other groups on visits 4 and 5. No other differences in questioning between the groups could be found.

The low-scoring group differed from the other groups in terms of follow-up questions and encouragement, which they received less across visits ($M_{\text{diff}} = 0.06$; $p < 0.05$). This was mostly caused by differences with the other groups at visit 4 ($M_{\text{diff}} = 0.07$; $p < 0.05$). In addition, a slightly higher overall proportion of follow-up questions and encouragements could be found for the high-scoring group compared to the mid-scoring group ($M_{\text{diff}} = 0.03$; $p < 0.05$), but the difference with the fluctuating group just fell short of significance ($M_{\text{diff}} = 0.03$; $p = 0.07$). When looking at individual visits, the proportion of follow-up questions and encouragement received by the high-scoring group differed significantly from that of the other groups at visit 2 ($M_{\text{diff}} = 0.15$; $p < 0.05$).

Table 4. Student utterances and proportion of answers and initiations across all visits, by group

	Mean number of utterances	Mean proportion of task-related answers	Mean proportion of student-initiated questions
High	71.03	0.87	0.02
Mid	102.65	0.71	0.07
Fluctuating	93.20	0.83	0.04
Low	86.96	0.69	0.06

Lastly, overall significant differences with regard to the proportion of explanations given by the teacher were found between the high-scoring group and the fluctuating group ($M_{\text{diff}} = -0.02$; $p < 0.05$) and between the high-scoring and the low-scoring groups ($M_{\text{diff}} = -0.02$; $p < 0.05$). This picture differed for the individual visits. At the first visit, the fluctuating group received more explanations compared to the high-scoring group ($M_{\text{diff}} = 0.04$; $p < 0.05$) and mid-scoring group ($M_{\text{diff}} = 0.03$; $p < 0.05$). At the second visit, the low-scoring group received more explanations than the high-scoring ($M_{\text{diff}} = 0.04$; $p < 0.05$) and mid-scoring groups ($M_{\text{diff}} = 0.03$; $p < 0.05$). However, during the third visit the groups received an equal proportion of explanations, and during the fourth visit the low-scoring group received fewer explanations than the high-scoring group ($M_{\text{diff}} = -0.02$; $p < 0.05$). The reverse could be observed during the fifth visit, when the low-scoring group again received more explanations compared to the high-scoring group ($M_{\text{diff}} = 0.04$; $p < 0.01$) and fluctuating group ($M_{\text{diff}} = 0.03$; $p < 0.05$).

Taking these observations together, the most consistent findings were fewer teacher utterances for the high-scoring group and more questioning for the low-scoring group, but more follow-up questions and encouragement for the high-scoring group. Differences in the proportion of explanations were more variable.

Student Behaviors. With regard to student behaviors (Table 4), the high-scoring group had a lower mean number of student utterances compared to the mid-scoring group ($M_{\text{diff}} = -31.62$; $p < 0.01$) and fluctuating group ($M_{\text{diff}} = -22.17$; $p < 0.05$), but not compared to the low-scoring group ($M_{\text{diff}} = -15.93$; $p = 0.08$). The mid- and low-scoring groups also differed significantly in their overall number of utterances ($M_{\text{diff}} = 15.69$; $p < 0.01$). These overall differences were mostly caused by smaller, non-significant differences on the individual visits, where only two significant differences could be found on the fourth visit, between the high-scoring and mid-scoring groups ($M_{\text{diff}} = -58$; $p < 0.05$), and between the mid- and low-scoring groups ($M_{\text{diff}} = 26.36$; $p < 0.05$). A final difference was found between the mid-scoring and the fluctuating groups on visit 2 ($M_{\text{diff}} = 35.42$; $p < 0.05$).

Despite the fact that the high-scoring group spoke less (i.e., had the lowest number of utterances), they did have a higher proportion of task-related answers compared to the mid-scoring group ($M_{\text{diff}} = 0.17$, $p < 0.01$) and the low-scoring group ($M_{\text{diff}} = 0.18$, $p < 0.01$). This pattern could be found on all individual visits ($p < 0.05$), except for the fifth one. The fluctuating group had a higher proportion of answers than the mid-scoring group ($M_{\text{diff}} = 0.12$; $p < 0.01$) and the low-scoring group ($M_{\text{diff}} = 0.14$, $p < 0.01$). The difference between the fluctuating and the mid-

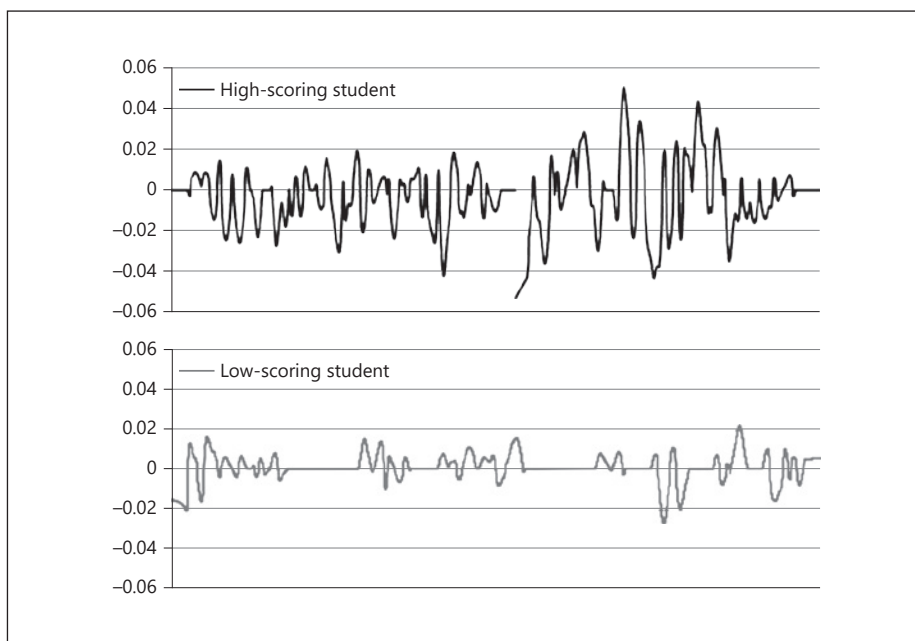


Fig. 4. Illustration of the teacher-student concordance values of a student each from the high-scoring and low-scoring groups during the first two visits. Positive values mean that the teacher's support and student's understanding levels show a similar change between two consecutive time points. Negative values indicate that the change in the teacher's support and student's understanding levels is dissimilar (see Methods section).

scoring group was visible at the second and third visits ($p < 0.05$), and the difference between the fluctuating and low-scoring groups was visible at the first and second visits ($p < 0.01$).

With regard to student-initiated questions, the only overall difference was found between the high-scoring and mid-scoring groups ($M_{\text{diff}} = -0.05$; $p < 0.01$). This difference was also significant for the first three individual visits ($p < 0.05$), but not for the final two. The mid-scoring group, with the highest proportion of student-initiated questions ($M = 0.07$), also differed significantly from the fluctuating and low-scoring groups during visit 3 ($M_{\text{diff}} = 0.09$ and 0.07 , respectively, $p < 0.01$). A final difference could be found between the high-scoring and low-scoring groups during visit 5, where the low-scoring group had more student-initiated questions ($M = 0.07$) than the high-scoring group ($M = 0.01$; $p < 0.05$).

To summarize, overall the high-scoring group had a lower number of student utterances, but a higher proportion of task-related answers. They had fewer student-initiated questions compared to the other groups. The mid-scoring group had the highest number of utterances and the highest proportion of student initiations, but fewer task-related answers. The low-scoring group had a higher number of utterances compared to the high-scoring group, but they also had the lowest proportion of task-related answers. They scored higher on student-initiated questions.

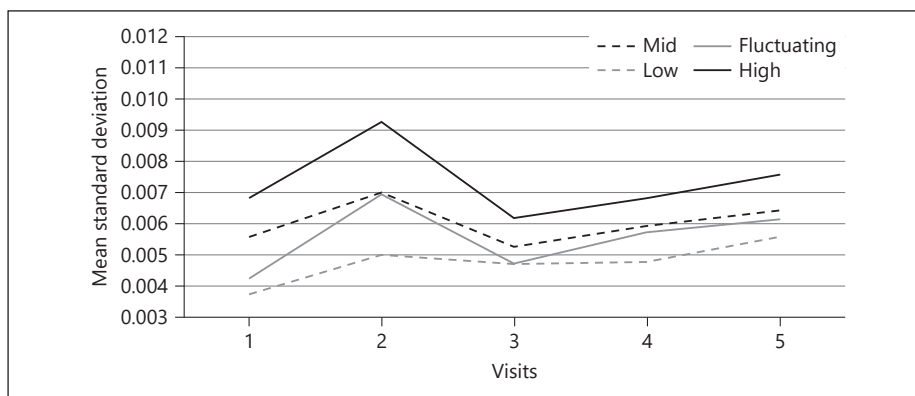


Fig. 5. Mean standard deviation of the positive teacher-student concordance values across the five visits, displayed by group.

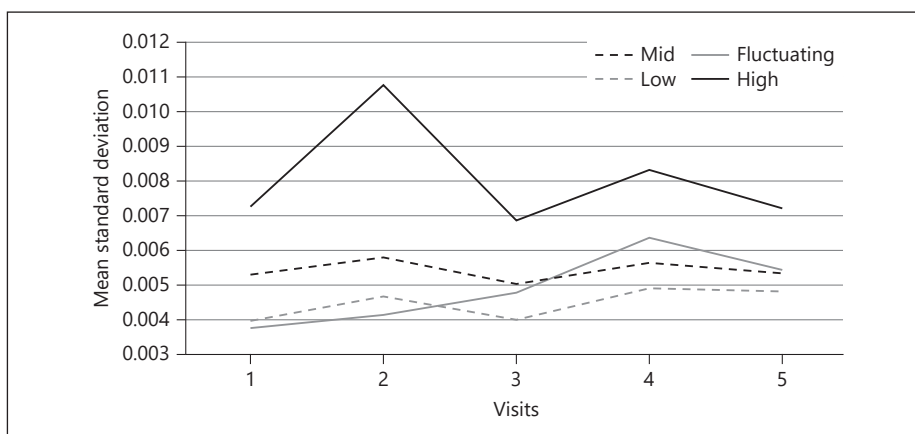


Fig. 6. Mean standard deviation of the negative teacher-student concordance values across the five visits, displayed by group.

Variability in Real-Time Interaction Patterns (RQ 2b)

Figure 4 shows an illustration of the teacher-student concordance of two cases, one from the high-scoring and one from the low-scoring group. Differences in the frequency and magnitude of concordance values immediately stand out. While the interaction of the low-scoring student has relatively few and small increases and decreases, the graph of the high-scoring student fluctuates considerably. This means that the interaction of the high-scoring student with the teacher is characterized by more moment-to-moment variability in teacher instructional support and student understanding. These variations consist of both similar changes (positive values, the

teacher's support and the student's understanding levels both increase or decrease), and dissimilar changes (negative values, either the support changes to a higher quality while the understanding level decreases, or the support changes to a lower quality while the understanding level increases).

To further investigate the moment-to-moment variability of the teacher's instructional support and the student's understanding levels, we calculated the standard deviation of the positive and negative concordance values for each visit and took the average for each group of students. The graphs that display the variation in concordance by group (Fig. 5, 6) show that the high-scoring group had the most variation in both positive and negative concordance values compared to the other groups. The low-scoring group showed the least variation in concordance values. This means that there are only few moment-to-moment similar and dissimilar changes in teacher support and student understanding. Generally, more variations can be found in the high-scoring and mid-scoring groups and the fewest variations in the low-scoring group.

Monte Carlo permutation tests revealed that both the positive and negative variations in concordance of the high-scoring group differed significantly from the lowest scoring group, on all visits ($p < 0.05$). The negative concordance variations of the high-scoring group also differed from the mid-scoring group on visits 2, 3, and 4 ($p < 0.05$), but not on visit 1 ($p = 0.12$) and visit 5 ($p = 0.06$). No difference in positive concordance variations could be found between the high-scoring and mid-scoring groups. A difference in positive concordance variations between the high-scoring and the fluctuating groups could be found at visits 1 and 3 ($p < 0.05$). The p values for the other visits were low but fell short of significance (between 0.07 and 0.13). A difference in negative concordance variations between the high-scoring and fluctuating groups was found on visits 1 and 2 ($p < 0.01$). The other visits again had low p values that just fell short of significance (between 0.06 and 0.08). The concordance values of the mid-scoring group did not differ from those of the fluctuating group. There were, however, some significant differences in positive concordance variations between the mid-scoring and the low-scoring groups, at visit 1 ($p < 0.05$), visit 2 ($p < 0.01$), and visit 4 ($p < 0.05$). Lastly, no significant differences between the low-scoring and fluctuating groups could be found.

To conclude, the data show that long-term developmental trajectories of students can be linked to the variability in concordance between teacher's support and student's understanding levels measured during their real-time interactions. Students with the highest understanding levels also had the most variation in teacher-student concordance. That is, their interactions were characterized by more moment-to-moment variations in the teacher's support and the student's level of understanding. The other three groups of students had overlapping concordance values, yet, in general, the mid-scoring and fluctuating groups had higher concordance values than the low-scoring group.

Discussion

The aim of this study was to couple learning trajectories of students' scientific understanding to short-term teacher-student dynamics. We investigated how 31 young students, from regular and special educational settings, developed their understanding of air pressure during five consecutive science tasks while they were guided

by a teacher, over the course of 18 months. We took a dynamic systems perspective, and used a scale based on skill theory (Fischer, 1980; Fischer & Bidell, 2006) to classify and conceptualize changes in students' understanding. Importantly, this metric is sensitive to fluctuations in understanding during single tasks but can also be used to detect changes in understanding on the long-term time scale (Fischer et al., 2007; Parziale & Fischer, 2002).

We first examined the long-term development of students' understanding over 18 months and investigated whether we could distinguish different groups of learners. Given that long-term development is assumed to depend on variations during and across learning tasks (Granott, 2002; Van Geert, 1997), which are partly due to changing contextual constraints (Fischer & Bidell, 2006; Yan & Fischer, 2002), we proceeded by examining differences between these groups in terms of teacher and student behaviors during each task and the variability within their real-time interaction patterns. In this way, the data of our study relate to Fogel and Garvey's (2007) idea of a "cooperative unit," and the study resembles Fischer and Granott's (1995) study on problem solving in dyads, and Steenbeek, Jansen, and Van Geert's (2012) longitudinal and microdevelopmental study of individual instruction sessions during arithmetic lessons.

Long-Term Developmental Trajectories of Understanding

An HAC analysis (Borgen & Barnett, 1987) yielded four groups of students with different learning trajectories. One of these groups was characterized as high-scoring, meaning that the complexity levels of their understanding exceeded the other three groups, on all visits. The mid-scoring group showed similar increases and decreases in understanding over time, albeit on a slightly lower level. The understanding levels of the low-scoring group remained relatively stable, after a small initial increase. Lastly, the fluctuating group showed considerable variations in understanding between visits, going from very low to very high complexity levels. A possible explanation for the considerable difference between the first and second visit may be found in Yan and Fischer (2002). According to these authors, students often show more extreme variations at the start of their learning process. Students in the fluctuating group had the lowest understanding levels at the start of the study (Fig. 3), which may indicate that these students had just started to build their understanding of the air pressure tasks.

Our quantitative and differential measures of the development of students' understanding of scientific phenomena show how different students of the same cohort can be in terms of learning. While this is not an unexpected finding, the shape of different learning trajectories rarely gets attention in research (see for some interesting qualitative examples: Löfgren & Helldén, 2009; Shapiro, 2010; Tytler & Peterson, 2004). Yet, knowing in what ways students can develop differently is important for researchers and educators alike, as it provides important clues on how to optimize these trajectories (Granott & Parziale, 2002), especially when they are coupled to microdevelopmental data of students' understanding during learning tasks (see below).

Interestingly, only a few differences in group composition with regard to age, gender, and school type were found. On average, the high-scoring group was signifi-

cantly older than the fluctuating and low-scoring groups, and its proportion of male students was significantly higher compared to the fluctuating group, but not compared to the other groups. The low-scoring group had significantly more students with emotional and behavioral disabilities than the fluctuating group, but not compared to the other two groups. The fact that only a few demographic differences could be found resonates with the theory of dynamic systems, and especially the premise that differences between students are to an important extent caused by factors that are intrinsic to the developmental process (Van Geert, 1998b). This means that regardless of age, gender, and school type, the development of students' understanding is the result of multiple self-organizing and changing components of both the student and the environment (Thelen & Corbetta, 2002). This may also be the reason why the students with emotional and behavioral disabilities, who generally display academic delays (Reid et al., 2004), could be found in all four clusters. Their development across these tasks depends not just on their academic status, but more so on their interaction with the teacher and the task material. A study of Van der Steen, Steenbeek, Wielinski, and Van Geert (2012b) also showed that these students are able to perform as well as students without emotional and behavioral problems when adequate instructional support is offered, tailored to the individual student. Hence, the crux of development is the person-context interaction that underlies it, not simply the student's characteristics. The current study therefore connected the different learning trajectories to concrete student and teacher behaviors during their interaction and the variability in their real-time interaction patterns.

Differences in Terms of Teacher Support and Student Behaviors

We found fewer teacher and student utterances for the high-scoring group and a smaller proportion of student-initiated questions, but a higher proportion of task-related answers. The teacher used relatively more questioning in interactions with students of the low-scoring group, but less follow-up questions and encouragement. The low-scoring group had a considerable number of student utterances compared to the other groups and the highest proportion of student-initiated questions, but also the lowest proportion of task-related answers. In sum, interactions in the high-scoring group were more compact and task-focused, while interactions in the low-scoring group were characterized by more questioning of both interaction partners, and fewer follow-up questions and encouragement.

While the literature reports that questioning, and especially open-ended questions, help students to construct their understanding (Meindertsma, van Dijk, Steenbeek, & van Geert, 2014; Topping & Trickey, 2014; Van Vondel et al., 2017), students in the low-scoring group did not seem to benefit from the higher proportion of questions compared to the other three groups. From a dynamic systems perspective, student and context form a continuous person-environment loop, which means that the student's understanding is the product of a dynamic interaction between person and context (Van der Steen et al., 2012a). At any moment in time, the student affects the teacher and vice versa, thereby creating the conditions under which they operate the next moment in time (Fogel & Garvey, 2007; Steenbeek, 2006). This means that it is unlikely that simply the teacher's behavior caused a difference in understanding levels between the groups. Rather, the interaction dynamics between student and teach-

er resulted in specific developmental patterns of understanding over time. The success of specific forms of teacher support, such as questioning, therefore highly depends on the student's reactions.

Differences in Real-Time Interaction Patterns

Researchers have argued that long-term development would evolve through a series of microdevelopmental sequences and would depend on variability in understanding during and across learning tasks (Granott, 2002; Van Geert, 1997). Variations in understanding are, however, embedded within the social environment (Brown et al., 1989; Lave & Wenger, 1991; Van der Steen et al., 2012a). We therefore investigated the coupling between the real-time variations in students' understanding and teacher's instructional support for the four groups of learners. Specifically, we calculated the concordance coefficient that indicates the similarity (positive values) or dissimilarity (negative values) of changes in the student's understanding and the teacher's support. Students from the high-scoring group had the most variability in teacher-student concordance. This means that more moment-to-moment changes in teacher instructional support and understanding levels were observed in the interactions of this group, both similar changes (e.g., higher-quality support – higher understanding levels) and dissimilar changes (e.g., higher-quality support – lower understanding levels). The low-scoring group had the lowest variability in concordance values and therefore fewer moment-to-moment changes in teacher instructional support and understanding levels.

Cognitive developmental patterns are often characterized by high variability (Yan & Fischer, 2002). In order to move forward to a higher level of understanding, the components of the system should be able to reassemble and break up old patterns (Siegler, 2007; Thelen & Corbetta, 2002). Interestingly, the highest scoring group in this study showed considerably more similar and dissimilar variations in teacher support and understanding levels compared to the other groups. Although the other groups had overlapping concordance values, in general the mid-scoring and fluctuating groups had higher concordance values than the low-scoring group. This study therefore provides important empirical evidence that variability is indeed the force that drives learning, as proposed by the theory of dynamic systems (Lichtwarck-Aschoff et al., 2009; Thelen et al., 1991; Van Geert & Van Dijk, 2002). Moreover, and importantly, this study shows that learning is not only associated with variability in students' abilities but is related to variability in the real-time interaction patterns *between* teacher and student. That is, the variability that is co-constructed by the teacher-student dyad, both when they change in a similar direction (e.g., higher-quality support – higher understanding levels) and when they change in an opposite direction (e.g., higher-quality support – lower understanding levels), is associated with different learning trajectories. Higher levels of understanding seem to occur in more “flexible” interactions.

Interestingly, similar findings have been found in clinical science. In a study on cognitive therapy for depression, variability in thought patterns, affect, and behavior was associated with more progress at the end of therapy (Hayes & Strauss, 1998). Furthermore, in a study on the effect of combined parent management training and cognitive-behavioral programs, flexibility during problem-solving

discussions in parent-child dyads was related to children's behavioral improvement at the end of the study (Granic, O'Hara, Pepler, & Lewis, 2007). Another study showed that dyadic flexibility in parent-child dyads at 3 years of age predicted lower levels of externalizing problems 2 and a half years later (Lunkenheimer, Olson, Hollenstein, Sameroff, & Winter, 2011). Yet, introducing variability in interactions has not become a mainstream strategy in clinical practice, nor is this the case in (special) educational settings. A possible explanation may be that variability in dyads is hard to observe, after all, it does not only refer to the student's behavior or the teacher's behavior, but to the combination of the two. When working together, teacher and student mutually affect each other and become a "cooperative unit" (Fogel & Garvey, 2007), with unique properties that transcend individual behavior. Future studies could investigate whether increased variability in teacher questions, strategies, or verbal complexity stimulates the mutual variability within the teacher-student dyad, thereby creating a more "flexible" interaction. This may be an especially important avenue for research on students with emotional and behavioral disabilities. After all, in this study, students with emotional and behavioral disabilities were scattered among the four different groups of learners, despite earlier reports of their significant academic delays (Reid et al., 2004).

Limitations and Future Directions

In this study, 31 students worked individually on five air pressure tasks while being guided by a teacher over the course of 18 months. In this way, our study can be considered a "hybrid approach" that connects a longitudinal view of the data with five periods of microdevelopment. Such an approach, that is, incorporating multiple timescales, may be useful for the study of change in science learning (Brock & Taber, 2017). The results of this study show four different groups of learners, which could be associated with specific student and teacher behaviors during the tasks, as well as with specific variations in their real-time interaction patterns. While an important aspect of dynamic systems is that real-time interaction patterns influence macrodevelopmental patterns, the opposite is also true. This circular relationship between micro- and macrodevelopment (see Introduction section) is not specifically addressed in this study. Our results show that specific long-term trajectories of the development of understanding are related to teacher-student dynamics during the science tasks, but the data do not reveal how macrodevelopmental patterns constrain the real-time interactions (Lewis, 2002).

It should also be noted that even though this study is based on coded video data of real-time teacher-student interactions, the quantitative computations make the link to the natural classroom situation less evident. Future research could focus on how interactions characterized by low variability can be detected in practice, and how teachers can stimulate such variability. In addition, although this study shows that more flexible teacher-student interactions could be found in the highest scoring group, we do not know whether these interactions can also become *too* flexible. Indeed, researchers have postulated that adaptive systems do show an optimal balance between flexibility and stability (Delignières & Marmelat, 2012; Den Hartigh, Cox, Gernigon, Van Yperen, & Van Geert, 2015; Goldberger et al., 2002;

Kello, Beltz, Holden, & Van Orden, 2007; Nourrit-Lucas, Tossa, Zelic, & Delignières, 2015; Wijnants, Bosman, Hasselman, Cox, & Van Orden, 2009; Wijnants, Hasselman, Cox, Bosman, & Van Orden, 2012). Lastly, the theory of dynamic systems suggests that some students may be sensitive to relatively small changes in components that are part of themselves or the environment, which may lead to a transition to a more advanced understanding (Thelen & Corbetta, 2002). An interesting avenue for future research would be to examine whether high-performing students are indeed more sensitive to small changes compared to students who perform less well.

Statement of Ethics

The parents or legal guardians of the students in this study have given their written informed consent. The study protocol has been approved by the research institute's committee on human research. The study was conducted ethically in accordance with the World Medical Association Declaration of Helsinki.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Source

This research was supported in part by a grant from the Dutch Ministry of Education (Curious Minds Program). The sponsor had no involvement in the study design, collection, analysis and interpretation of the data, writing of the article, and the decision to submit the article for publication.

Author Contributions

S.S. collected data, conceived and designed the analyses, performed the analyses, and wrote the article. H.W.S. wrote the article and provided a critical revision of the article. R.J.R.H. wrote the article, contributed to the theoretical framework, and provided a critical revision of the article. P.L.C.G. conceived and designed the analyses, contributed analysis tools, and provided a critical revision of the article. The manuscript has been seen and approved by all authors.

References

- Borgen, F. H., & Barnett, D. C. (1987). Applying cluster analysis in counseling psychology research. *Journal of Counseling Psychology, 34*(4), 456–468. <https://doi.org/10.1037/0022-0167.34.4.456>
- Bos, J., & Steenbeek, H. (2006). *Mediacoder: A simple application for coding behavior within media files*. Groningen, The Netherlands: University of Groningen.
- Brock, R., & Taber, K. S. (2017). The application of the microgenetic method to studies of learning in science education: Characteristics of published studies, methodological issues and recommendations for future research. *Studies in Science Education, 53*(1), 45–73. <https://doi.org/10.1080/03057267.2016.1262046>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42. <https://doi.org/10.3102/0013189X018001032>

- Clement, J. J. (2009). The role of imagistic simulation in scientific thought experiments. *Topics in Cognitive Science*, 1(4), 686–710. <https://doi.org/10.1111/j.1756-8765.2009.01031.x>
- Delignières, D., & Marmelat, V. (2012). Fractal fluctuations and complexity: Current debates and future challenges. *Critical Reviews in Biomedical Engineering*, 40(6), 485–500. <https://doi.org/10.1615/Crit-RevBiomedEng.2013006727>
- Den Hartigh, R. J. R., Cox, R. F., Gernigon, C., Van Yperen, N. W., & Van Geert, P. L. (2015). Pink noise in rowing ergometer performance and the role of skill level. *Motor Control*, 19(4), 355–369. <https://doi.org/10.1123/mc.2014-0071>
- Den Hartigh, R. J., Van Der Steen, S., De Meij, M., Van Yperen, N. W., Gernigon, C., & Van Geert, P. L. (2014). Characterising expert representations during real-time action: A Skill Theory application to soccer. *Journal of Cognitive Psychology*, 26(7), 754–767. <https://doi.org/10.1080/20445911.2014.955028>
- Den Hartigh, R. J. R., Van Der Steen, S., Hakvoort, B., Frencken, W. G. P., & Lemmink, K. A. P. M. (2018). Differences in game reading between selected and non-selected youth soccer players. *Journal of Sports Sciences*, 36(4), 422–428. <https://doi.org/10.1080/02640414.2017.1313442>
- Den Hartigh, R. J., Van Dijk, M. W., Steenbeek, H. W., & Van Geert, P. L. (2016). A dynamic network model to explain the development of excellent human performance. *Frontiers in Psychology*, 7, 532. <https://doi.org/10.3389/fpsyg.2016.00532>
- Dunlap, G., Strain, P. S., Fox, L., Carta, J. J., Conroy, M., Smith, B. J., . . . Sowell, C. (2006). Prevention and intervention with young children's challenging behavior: Perspectives regarding current knowledge. *Behavioral Disorders*, 32(1), 29–45. <https://doi.org/10.1177/019874290603200103>
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477–531. <https://doi.org/10.1037/0033-295X.87.6.477>
- Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In R. M. Lerner (Ed.) & W. Damon (Series Ed.), *Handbook of child psychology: Vol. I. Theoretical models of human development* (pp. 313–399). New York, NY: Wiley.
- Fischer, K. W., & Granott, N. (1995). Beyond one-dimensional change: Parallel, concurrent, socially distributed processes in learning and development. *Human Development*, 38(6), 302–314. <https://doi.org/10.1159/000278336>
- Fischer, K. W., Rose, L. T., & Rose, S. P. (2007). Growth cycles of mind and brain: Analyzing developmental pathways of learning disorders. In K. W. Fischer, J. Holmes Bernstein, & M. H. Immordino-Yang (Eds.), *Mind, brain, and education in reading disorders* (pp. 101–123). New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511489952.007>
- Fischer, K. W., & Van Geert, P. (2014). Dynamic development of brain and behavior. In P. C. M. Molenaar, R. M. Lerner, & K. M. Newell (Eds.), *Handbook of developmental systems theory and methodology* (pp. 287–315). New York, NY: Guilford Press.
- Flynn, E., & Siegler, R. (2007). Measuring change: Current trends and future directions in microgenetic research. *Infant and Child Development*, 16(1), 135–149. <https://doi.org/10.1002/icd.502>
- Fogel, A., & Garvey, A. (2007). Alive communication. *Infant Behavior and Development*, 30(2), 251–257. <https://doi.org/10.1016/j.infbeh.2007.02.007>
- Gelman, R., Romo, L., & Francis, W. S. (2002). Notebooks as windows on learning: The case of a science-into-ESL program. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 269–293). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.011>
- Goldberger, A. L., Amaral, L. A., Hausdorff, J. M., Ivanov, P. C., Peng, C. K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 1), 2466–2472. <https://doi.org/10.1073/pnas.012579499>
- Granic, I., O'Hara, A., Pepler, D., & Lewis, M. D. (2007). A dynamic systems analysis of parent-child changes associated with successful “real-world” interventions for aggressive children. *Journal of Abnormal Child Psychology*, 35(5), 845–857. <https://doi.org/10.1007/s10802-007-9133-4>
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review*, 113(1), 101–131. <https://doi.org/10.1037/0033-295X.113.1.101>
- Granott, N. (1998). We learn, therefore we develop: Learning versus development – or developing learning? In M. C. Smith & T. Pourchot (Eds.), *The educational psychology series. Adult learning and development: Perspectives from educational psychology* (pp. 15–34). Mahwah, NJ: Erlbaum Associates.
- Granott, N. (2002). How microdevelopment creates macrodevelopment: Reiterated sequences, backward transitions, and the zone of current development. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 213–242). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.009>

- Granott, N., Fischer, K. W., & Parziale, J. (2002). Bridging to the unknown: A transition mechanism in learning and development. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 131–156). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.006>
- Granott, N., & Parziale, J. (2002). Microdevelopment: A process-oriented perspective for studying development and learning. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 1–28). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.001>
- Guevara, M., Cox, R. F. A., van Dijk, M., & van Geert, P. (2017). Attractor dynamics of dyadic interaction: A recurrence based analysis. *Nonlinear Dynamics Psychology and Life Sciences*, *21*(3), 289–317.
- Guevara, M., van Dijk, M., & van Geert, P. (2016). Microdevelopment of peer interactions and scientific reasoning in young children/Microdesarrollo de la interacción entre pares y el razonamiento científico en niños pequeños. *Infancia y Aprendizaje*, *39*(4), 727–771. <https://doi.org/10.1080/02103702.2016.1215083>
- Halley, J. D., & Winkler, D. A. (2008). Classification of emergence and its relation to self organization. *Complexity*, *13*(5), 10–15. <https://doi.org/10.1002/cplx.20216>
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*(5), 949–967. <https://doi.org/10.1111/j.1467-8624.2005.00889.x>
- Hayes, A. M., & Strauss, J. L. (1998). Dynamic systems theory as a paradigm for the study of change in psychotherapy: An application to cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, *66*(6), 939–947. <https://doi.org/10.1037/0022-006X.66.6.939>
- Howe, M. L., & Lewis, M. D. (2005). The importance of dynamic systems approaches for understanding development. *Developmental Review*, *25*(3-4), 247–251. <https://doi.org/10.1016/j.dr.2005.09.002>
- Jacoby, W. G. (2000). Loess: A nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*, *19*(4), 577–613. [https://doi.org/10.1016/S0261-3794\(99\)00028-1](https://doi.org/10.1016/S0261-3794(99)00028-1)
- Kello, C. T., Beltz, B. C., Holden, J. G., & Van Orden, G. C. (2007). The emergent coordination of cognitive function. *Journal of Experimental Psychology: General*, *136*(4), 551–568. <https://doi.org/10.1037/0096-3445.136.4.551>
- Lane, K. L., Barton-Arwood, S. M., Nelson, J. R., & Wehby, J. (2008). Academic performance of students with emotional and behavioral disorders served in a self-contained setting. *Journal of Behavioral Education*, *17*(1), 43–62. <https://doi.org/10.1007/s10864-007-9050-1>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>
- Lee, K., & Karmiloff-Smith, A. (2002). Macro-and microdevelopmental research: Assumptions, research strategies, constraints, and utilities. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 243–265). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.010>
- Lewis, M. D. (2002). Interacting time scales in personality (and cognitive) development: Intentions, emotions, and emergent forms. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 183–212). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.008>
- Lichtwarck-Aschoff, A., Kunnen, S. E., & van Geert, P. L. (2009). Here we go again: A dynamic systems perspective on emotional rigidity across parent-adolescent conflicts. *Developmental Psychology*, *45*(5), 1364–1375. <https://doi.org/10.1037/a0016713>
- Lichtwarck-Aschoff, A., van Geert, P., Bosma, H., & Kunnen, S. (2008). Time and identity: A framework for research and theory formation. *Developmental Review*, *28*(3), 370–400. <https://doi.org/10.1016/j.dr.2008.04.001>
- Löfgren, L., & Helldén, G. (2009). A longitudinal study showing how students use a molecule concept when explaining everyday situations. *International Journal of Science Education*, *31*(12), 1631–1655. doi:10.1080/09500690802154850
- Lunkenheimer, E. S., Olson, S. L., Hollenstein, T., Sameroff, A. J., & Winter, C. (2011). Dyadic flexibility and positive affect in parent-child coregulation and the development of child behavior problems. *Development and Psychopathology*, *23*(2), 577–591. <https://doi.org/10.1017/S095457941100006X>
- Meindertsma, H. B., van Dijk, M. W., Steenbeek, H. W., & van Geert, P. L. (2014). Assessment of preschooler's scientific reasoning in adult-child interactions: What is the optimal context? *Research in Science Education*, *44*, 215–237. <https://doi.org/10.1007/s11165-013-9380-z>

- Montes, J., Van Dijk, M., Puche, R., & Van Geert, P. (2017). Trajectories of scientific reasoning: A micro-genetic study on children's inquiry functioning. *Journal for Person-Oriented Research*, *3*(2), 67–85. <https://doi.org/10.17505/jpor.2017.07>
- Newell, K. M., Liu, Y. T., & Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychological Review*, *108*(1), 57–82. <https://doi.org/10.1037/0033-295X.108.1.57>
- Nourrit-Lucas, D., Tossa, A. O., Zélic, G., & Delignières, D. (2015). Learning, motor skill, and long-range correlations. *Journal of Motor Behavior*, *47*(3), 182–189. <https://doi.org/10.1080/00222895.2014.967655>
- Parziale, J., & Fischer, K. W. (1998). The practical use of skill theory in classrooms. In R. J. Sternberg & W. M. Williams (Eds.), *Intelligence, instruction, and assessment* (pp. 95–110). Mahwah, NJ: Erlbaum Associates.
- Piaget, J. (2001). *The psychology of intelligence* (rev. ed.) (M. Piercy & D. E. Belyne, Trans.). London, UK: Routledge Classics (original work published 1947).
- Rakotomalala, R. (2005). Tanagra: A free software for research and academic purposes. In *Proceedings of European Geothermal Congress* (pp. 697–702).
- Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education*, *38*(3), 130–143. <https://doi.org/10.1177/00224669040380030101>
- Rogoff, B., & Lave, J. (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Schwartz, M. S., & Fischer, K. W. (2004). Building general knowledge and skill: Cognition and micro-development in science learning. In A. Demetriou & A. Raftopoulos (Eds.), *Emergence and transformation in the mind: Modeling and measuring cognitive change* (pp. 157–185). Cambridge, UK: Cambridge University Press.
- Shapiro, B. (2010). Studying lifeworlds of science learning: A longitudinal study of changing ideas, contexts, and personal orientations in science learning. *Canadian Journal of Mathematics Science and Technology Education*, *4*, 127–147. <https://doi.org/10.1080/14926150409556601>
- Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, *28*(3), 225–273. <https://doi.org/10.1006/cogp.1995.1006>
- Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. <https://doi.org/10.1111/j.1467-7687.2007.00571.x>
- Steenbeek, H. W. (2006). *Modeling dyadic child-peer interactions: Sociometric status, emotional expressions, and instrumental actions during play* (Unpublished doctoral dissertation), University of Groningen, The Netherlands.
- Steenbeek, H., Jansen, L., & van Geert, P. (2012). Scaffolding dynamics and the emergence of problematic learning trajectories. *Learning and Individual Differences*, *22*(1), 64–75. <https://doi.org/10.1016/j.lindif.2011.11.014>
- Steenbeek, H., & van Geert, P. (2013). The emergence of learning-teaching trajectories in education: A complex dynamic systems approach. *Nonlinear Dynamics Psychology and Life Sciences*, *17*(2), 233–267.
- Steenbeek, H., van Geert, P., & van der Steen, S. (2017). Using a complex dynamic systems approach to interaction research in education: The case of autism interventions. In M. S. Schwartz & E. J. Paré-Blagoev (Eds.), *Research in mind, brain, and education* (pp. 182–207). London, UK: Routledge. <https://doi.org/10.4324/9781315670621-8>
- Thelen, E., & Corbetta, D. (2002). Microdevelopment and dynamic systems: Applications to infant motor development. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 59–79). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.003>
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press.
- Thelen, E., Ulrich, B. D., & Wolff, P. (1991). Hidden skills: A dynamic systems analysis of treadmill stepping during the first year. *Monographs of the Society for Research in Child Development*, *56*(1), 1–98. <https://doi.org/10.2307/1166099>
- Todman, J. B., & Dugard, P. (2001). *Single-case and small-N experimental designs: A practical guide to randomization tests*. Mahwah, NJ: Erlbaum Associates. <https://doi.org/10.4324/9781410600943>

- Topping, K. J., & Trickey, S. (2014). The role of dialog in philosophy for children. *International Journal of Educational Research*, *63*, 69–78. <https://doi.org/10.1016/j.ijer.2013.01.002>
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*(1), 7–26. [https://doi.org/10.1016/S0022-4405\(01\)00092-9](https://doi.org/10.1016/S0022-4405(01)00092-9)
- Turvey, M. T. (2009). On the notion and implications of organism-environment system. *Ecological Psychology*, *21*(2), 97–111. <https://doi.org/10.1080/10407410902877041>
- Turvey, M. T., & Shaw, R. E. (1999). Ecological foundations of cognition: Symmetry and specificity of animal-environment systems. In R. Nunez & W. J. Freeman (Eds.), *Reclaiming cognition: The primacy of action, intention and emotion* (pp. 95–110). Thorveston, UK: Imprint Academic. <https://doi.org/10.1080/00071773.2002.11007363>
- Tytler, R. (1998). The nature of students' informal science conceptions. *International Journal of Science Education*, *20*(8), 901–927. <https://doi.org/10.1080/0950069980200802>
- Tytler, R., & Peterson, S. (2004). Young children learning about evaporation: A longitudinal perspective. *Canadian Journal of Mathematics Science and Technology Education*, *4*, 111–126. <https://doi.org/10.1080/14926150409556600>
- Van Der Steen, S., Steenbeek, H., Van Dijk, M., & Van Geert, P. (2014). A process approach to children's understanding of scientific concepts: A longitudinal case study. *Learning and Individual Differences*, *30*, 84–91. <https://doi.org/10.1016/j.lindif.2013.12.004>
- Van der Steen, S., Steenbeek, H., & Van Geert, P. (2012a). Using the dynamics of a person context system to describe children's understanding of air pressure. In H. Kloos, B. J. Morris, & J. L. Amaral (Eds.), *Current topics in children's learning and cognition*. Rijeka, Croatia: IntechOpen. <https://doi.org/10.5772/53935>
- Van Der Steen, S., Steenbeek, H., Wielinski, J., & Van Geert, P. (2012b). A comparison between young students with and without special needs on their understanding of scientific concepts. *Education Research International*, *2012*, 1–12. <https://doi.org/10.1155/2012/260403>
- Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*(1), 3–53. <https://doi.org/10.1037/0033-295X.98.1.3>
- Van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. Hertfordshire, England: Harvester Wheatsheaf.
- Van Geert, P. (1995). Dimensions of change: A semantic and mathematical analysis of learning and development. *Human Development*, *38*(6), 322–331. <https://doi.org/10.1159/000278338>
- Van Geert, P. (1998a). A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky, and beyond. *Psychological Review*, *105*(4), 634–677. <https://doi.org/10.1037/0033-295X.105.4.634-677>
- Van Geert, P. (1998b). We almost had a great future behind us: The contribution of non-linear dynamics to developmental-science-in-the-making. *Developmental Science*, *1*(1), 143–159. <https://doi.org/10.1111/1467-7687.00020>
- Van Geert, P. (2002). Developmental dynamics, intentional action, and fuzzy sets. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 319–342). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.013>
- Van Geert, P. L. C. (1997). Time and theory in social psychology. *Psychological Inquiry*, *8*(2), 143–151. https://doi.org/10.1207/s15327965pli0802_11
- Van Geert, P., & Fischer, K. W. (2009). Dynamic systems and the quest for individual-based models of change and development. In J. P. Spencer, M. S. C. Thomas, & J. McClelland (Eds.), *Toward a new grand theory of development? Connectionism and dynamic systems theory reconsidered*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195300598.003.0016>
- Van Geert, P., & Steenbeek, H. (2005a). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, *25*(3-4), 408–442. <https://doi.org/10.1016/j.dr.2005.10.003>
- Van Geert, P., & Steenbeek, H. (2005b). The dynamics of scaffolding. *New Ideas in Psychology*, *23*(3), 115–128. <https://doi.org/10.1016/j.newideapsych.2006.05.003>
- Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, *25*(4), 340–374. [https://doi.org/10.1016/S0163-6383\(02\)00140-6](https://doi.org/10.1016/S0163-6383(02)00140-6)
- Van Vondel, S., Steenbeek, H., van Dijk, M., & van Geert, P. (2017). The effects of video feedback coaching for teachers on scientific knowledge of primary students. *Research in Science Education*, *48*(2), 301–324. <https://doi.org/10.1007/s11165-016-9569-z>

- Vandenbroucke, L., Spilt, J., Verschueren, K., Piccinin, C., & Baeyens, D. (2018). The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher-student interactions for children's executive functions. *Review of Educational Research, 88*(1), 125–164. <https://doi.org/10.3102/0034654317743200>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, MA: Harvard University Press.
- Wehby, J. H., Lane, K. L., & Falk, K. B. (2003). Academic instruction for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 11*(4), 194–197. <https://doi.org/10.1177/10634266030110040101>
- Wijnants, M. L., Bosman, A. M., Hasselman, F., Cox, R. F., & Van Orden, G. C. (2009). 1/f scaling in movement time changes with practice in precision aiming. *Nonlinear Dynamics Psychology and Life Sciences, 13*(1), 79–98.
- Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. (2012). An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia, 62*(2), 100–119. <https://doi.org/10.1007/s11881-012-0067-3>
- Yan, Z., & Fischer, K. (2002). Always under construction. *Human Development, 45*(3), 141–160. <https://doi.org/10.1159/000057070>