University of Groningen

A complexity approach to investigating the effectiveness of an intervention for lower grade teachers on teaching science

Wetzels, Anna; Steenbeek, Henderien; van Geert, Paul

Published in:
Complicity: An International Journal of Complexity and Education

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Citation for published version (APA):
Wetzels, A., Steenbeek, H., & van Geert, P. (2016). A complexity approach to investigating the effectiveness of an intervention for lower grade teachers on teaching science. Complicity: An International Journal of Complexity and Education, 13(1), 81-104.

A Complexity Approach to Investigating the Effectiveness of an Intervention for Lower Grade Teachers on Teaching Science

ANNEMIE WETZELS, HENDERIEN STEENBEEK, AND PAUL VAN GEERT
University of Groningen (The Netherlands)

This article describes the effectiveness and sustainability of teacher professional development interventions from a complexity viewpoint as well as a more 'standard' viewpoint. The first aim of this study is to give a theoretical overview of effective aspects of interventions regarding teachers' professionalizing using recent literature. The second aim is to re-interpret effectiveness and effectiveness studies using a complexity approach. The third aim is to empirically illustrate a complexity approach to the effectiveness of interventions using a multiple case study. We have described intervention specific aspects, teacher specific aspects, context specific aspects and implementation specific aspects and have shown in the cases that during an intervention these aspects intertwine and act as a complex dynamic system. The complexity approach to interventions has implications for future empirical research as well as for the design of interventions. In future research, results of small scale and large scale research should be combined, in order to obtain a better insight in all relevant aspects and their effect on teachers' behavior. Research ought to concentrate on the effects that all contextual aspects have intertwiningly in order to design more effective interventions and with better implementation processes.

Introduction

Children who are currently in preschool will enter the labor force around 2030, and will work until their retirement, around 2070. Nobody knows what the world will look like then, but it is likely that in order for children to be prepared for their professional career in the future, they must learn the skills for the 21st-century, namely critical thinking, problem solving, and decision making, among other skills (Binkley et al., 2012). The recognition that teachers play a central role in children's mastering of these skills is a necessary premise to acknowledge when considering interventions in education systems aimed at preparing children for the rapidly changing world (Ferguson, 1991; Grigg, Kelly, Gamoran, & Borman, 2012; Harwell, 2003; Hattie,

2003). If society aims to prepare pupils as well as possible for the future, teachers need to teach their pupils the latest developments in numerous areas. This also implies that, in order to keep pace with a rapidly changing world, teachers' professional development needs to continue after their formal education. However, much of the professional development that the market offers to teachers could be more effective than it is now (Garet et al., 2011; Gersten, Taylor, Keys, Rolfhus, & Newman-Gonchar, 2014; Mayer, 2004; Pressley, Graham, & Harris, 2006). The reason for this lack of effectiveness is not always poor design of the intervention itself, but rather the multitude of intertwining factors in the complex system surrounding a teacher and his [sic] classroom that influence the effectiveness of an intervention (Goldspink, 2007). Moreover, an intervention can be effective during the intervention period, but the question is whether these positive effects are sustainable when the interveners and the support and attention they give to school and teachers are no longer in place. Often the effects disappear relatively soon after the intervention is carried out (Han & Weiss, 2005; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997).

In this paper we describe the effectiveness and sustainability of teacher professional development interventions from a complexity viewpoint as well as from a more 'standard' viewpoint. The first aim of this study is to give a theoretical overview of effective aspects of interventions regarding teachers' professionalizing using recent literature. The second aim is to re-interpret effectiveness and effectiveness studies using a complexity approach. The third aim is to empirically illustrate a complexity approach to the effectiveness of interventions using a multiple case study (Hetherington, 2013).

2 Explaining the Effectiveness of Interventions Aimed at Teachers' Professional Development

2.1 Professional development

Professional development for teachers is a continuous process of professional growth (Day, 1999; Evans, 2002; Keiny, 1994) and encompasses all types of facilitated learning opportunities as well as informal learning opportunities situated in practice (Day, 1999; Lieberman, 1996). It is a necessary process because, first, it enables teachers to learn to use the knowledge acquired in teacher training in practice, and, after that, it helps to keep up with developments in society. After initial teacher training, the first part of professional development consists of the experience a new teacher gains at his or her job. As early as 1938, Dewey observed that this increase in experience alone did not necessarily result in more sustained knowledge or sustained better teaching. Unsurprisingly, the learning curve of teachers rapidly declines, as teachers' learning stagnates after the first few years working in practice (Barth, 2004). To keep the learning curve constant, teachers are provided with an overwhelming choice in professional development activities (Hattie, 2009). However, it is likely that most training opportunities are not effective in creating sustained improvement of teacher effectiveness because they are not in line with teachers' motivational and intellectual requirements (Day, 1999; Imperley, Wilson, Barrar, & Fung, 2007).

The criteria for the effectiveness of a professional development intervention for teachers are ambiguous (Guskey, 2003). For this paper, we use a classification applied by Garet, Porter, Desimone, Birman, & Yoon (2001) and Desimone (2009). According to this classification, teachers' change in knowledge and attitudes, teachers' change in behavior in the classroom, and pupils' results are the three elements that provide indications for the intervention's effectiveness (Desimone, 2009; Garet et al., 2001). Achieving these goals in the long run, and thus providing an effective and sustainable intervention, depends not only on the intervention itself but also on whether the intervention has been successfully implemented (Durlak & DuPre, 2008; Harwell, 2003; Noell & Gansle, 2009). Unfortunately, the effects of interventions tend to disappear relatively soon after the intervention is carried out (Han & Weiss, 2005; Noell et al., 1997). What, in addition to the quality of implementation, does the available literature tell us about the main aspects that have an influence on the effectiveness of professional development interventions?

2.2 Important aspects of a professional development intervention

Teachers are highly critical consumers of professional development activities. Day (1999) states that professional development for teachers "is not simply a cognitive process … it demands emotional commitment. It will involve the head and the heart" (p. 47). Garet et al. (2001) and Desimone et al. (2009) make a distinction between various aspects that positively contribute to the effectiveness of an intervention. The first concerns the properties of the intervention itself, the intervention-specific aspects. The second includes those variables relate to the teacher as a person, and the third relates to the school organization within which a teacher works, i.e., the school context (Han & Weiss, 2005). In addition to these variables, we must distinguish aspects of implementation quality.

2.2.1 Intervention specific aspects

Interventions are characterized by three structural and three substantive characteristics (Armour & Yelling, 2004; Birman, Desimone, Porter, & Garet, 2000; Garet et al., 2001). Structural elements are form, duration and collective participation (Garet et al., 2001). As for form, traditional forms of education, such as workshops and external trainings without follow-up in the classroom, are not as effective as training within the school or classroom with a connection to the teaching practice (Fabiano et al., 2013; Goldenberg & Gallimore, 1991; Loucks-Horsley, Stiles, Mundry, Love, & Hewson, 2010). Teacher training in combination with direct performance feedback produces the strongest training effect (Noell, et. al.; Rose & Church, 1998). The duration aspect can be understood in two ways: the period of time over which the intervention is spread, and the number of contact hours. Literature suggests that both components of duration need to be long enough in order to be effective in passing on knowledge and in sustaining changes over time (Garet et al., 2001; Harwell, 2003), although schools often do not have enough time or money to undertake longer trainings. Collective participation, i.e., participation by a group of teachers from the same school, is also a positive structural aspect of interventions that allows teachers who work together to share learning experiences, so they can integrate their new learning experiences with their own school background and curriculum (Garet et al., 2001; Harwell, 2003).

As regards the substantive elements in an intervention, first, there must be a focus on content. Teachers often lack subject matter knowledge, including the related teaching practices in specific content domains. Together these are known as the pedagogical content knowledge (Shulman, 1987). When comparing interventions based on general pedagogical knowledge, or on content knowledge, or on pedagogical content knowledge, the literature shows more positive effects on students achievements when the education activities are focused on the pedagogical content knowledge (Cohen & Hill, 2000; Loucks-Horsley et al., 2010; Smith, 1999), apparently because the latter gives teachers more concrete ideas for their lessons. Secondly, the intervention needs to be based on active learning, by means of reflection on the way a subject can be taught in the classroom, such as discussion and dialogue. In this way, the participating teachers obtain opportunities to practice what they learn in the classroom and discuss it with a colleague or a coach (Harwell, 2003). Finally, coherence with other learning activities is necessary (Garet et al., 2001) so that something learned can be integrated with a range of activities within a school.

2.2.2 Teacher specific aspects

To begin, teachers' sense of self-efficacy is important for their level of openness to new ideas (Berman, McLaughlin, Bass, Pauly, & Zallman, 1977; Durlak & DuPre, 2008). In addition, teachers should feel the need for using the learned content in their own classroom (Harwell, 2003), and they should feel intrinsically motivated to use the new practices in their daily routines because teachers' intrinsic motivation is very often the basis from which they started teaching (Dinham & Scott, 2004). Humans in general are basically in search for self-motivation and self-development since these aspects contribute to human well-being (Deci & Ryan, 2002; Diener, 1984).

When performing a school intervention, it must be kept in mind that attempts toward changing teachers' habits and behaviors can be a threat to teachers' intrinsic motivation in the classroom. More specifically, it is important that teachers' basic concerns (Frijda, 1986), namely autonomy, relatedness and competence (Minnaert, Boekaerts, & de Brabander, 2007; Ryan & Deci, 2000; Steenbeek & Van Geert, 2007) should not be threatened. Autonomy plays a role in that teachers should feel like they can make their own decisions concerning when an intervention is carried out at a school. Secondly, teachers often feel like they lack the competence to act the way they are expected to during the intervention (Goodrum, Hackling, & Rennie, 2001; Jarvis & Pell, 2004; Palmer, 2002). Primary teachers may take a general degree in teaching; consequently they may not have content knowledge of specific domains such as science, which may result in feeling uncomfortable during professionalization courses aimed at teaching science to young children, for instance. And finally, teachers need to feel connected with colleagues and pupils when teaching, especially in a period when things change in their classroom.

2.2.3 Context specific aspects

Changes in teachers' behavior do not occur in the vacuum of the classroom only. Various political, bureaucratic and systemic factors influence teachers' adaptation to new working methods in schools (Adelman & Taylor, 2003; Goldspink, 2007). In this study, we shall confine ourselves to the school specific factors because these factors can be addressed directly during an

intervention. They include organizational support in training, for instance, extra time for study, practice and reflection on each other's practices with co-workers, as well as supervision and support by the school principal during the intervention's implementation and afterwards (Fullan, 2007; Gottfredson & Gottfredson, 2002). The school system should be involved in a way that, once external support from trainers or researchers ends, the system is ready to integrate the elements of the intervention into the existing practices. All levels of an organization, i.e., the classroom level, the teacher team level, and the level of the school board should work together in order to create sustainable change (Coburn, 2003; Goldspink, 2007). Shared decision making regarding the intervention is an aspect that adds to a better implementation (Durlak & DuPre, 2008). If the intervention focuses on a group of teachers from the same school, it can promote teachers to work together, share mutual experiences, and find solutions to joint difficulties in the new way of working (Loucks-Horsley et al., 2010).

2.2.4 Aspects regarding implementation

As Boomstra (2014) describes, implementation (i.e., realizing the intervention) is an important aspect to consider when assessing the effectiveness of an intervention. The theoretical basis of a an intervention can be adequate, but the effectiveness can be negatively influenced by less sufficient execution of the intervention, or even by intervention activities that were never prescribed. Gearing et al. (2011) described four elements that contribute to minimizing the distance between the theoretical basis of an intervention and the delivery of components of the intervention in real practice: 1) good intervention design and protocols, 2) good intervention training, 3) sufficient monitoring of intervention delivery, and 4) sufficient monitoring of intervention receipt. These four components also need to be taken into account when designing an effective intervention. In addition, treatment integrity is an important aspect that greatly contributes to successful interventions. Treatment integrity is the extent to which an intervention is implemented as it was designed. When an intervention is not provided by trainers as intended, whatever effects occur will not be due to the intervention as such (Noell, Gresham, & Gansle, 2002). Documentation of the intervention, for example, can assure treatment adherence, whereas training of the trainers according to the documentation can assure their competence in providing the intervention as intended (Perepletchikova & Kazdin, 2005).

In fact, models that describe the effectiveness and sustainability of teacher professional development interventions from a standard viewpoint assume that a linear addition of all these factors can explain the effectiveness of an intervention in general. This might be true if a model is viewed as a purely statistical statement. That is to say, if the individual and temporal aspects of the intervention process are not taken into account, the effectiveness of an intervention across contexts and occasions can indeed be "explained" on the basis of how it combines and implements the elements discussed in the preceding section. However, the actual performing of an intervention in a concrete context cannot be modelled as a sum of the variables (Molenaar & Campbell, 2009). First of all, the actual performing of an intervention in a concrete context is a process, taking place in time. Second, it is a process, taking place in a particular context, with specific individuals and people who carry out the treatment. Third, it is a process taking place in a complex system, namely, a system of intertwining teacher, child, school, cultural and

community variables. It is highly likely that the statistical model emerging from sample-based effectiveness studies – which is a linear additive model — does not apply to the individual and complex processes of performing particular interventions. The point discussed in the next section is that understanding intervention effectiveness requires a complexity point of view in order to capture the temporal and idiosyncratic aspects of interventions. Nonetheless, it is important to make a distinction between designing an intervention, and performing an intervention. While designing an intervention, one can use the standard statistical model of effective components as a partial guideline. However, this is only a partial guideline in the sense that the designer of the intervention should be aware that performing the intervention will take place in a complex dynamic system with many intertwined and dynamically changing processes that are likely to be highly idiosyncratic, and in which the researcher inevitably becomes an embedded component, even if the researcher is not carrying out the intervention him or herself (Van Geert & Steenbeek, 2014). In the next section, we shall try to redefine teachers' professional development using a complexity approach.

2.3 Teachers' professional development: a complexity approach

Looking at professional development as a simple cause-and-effect process does not do justice to the complexity of functioning as a teacher in a specific school context, and to the complex world in which teachers function (Goldspink, 2007; Opfer & Pedder, 2011). Complexity research reasons that various "forces" within a person and within society act simultaneously and mutually to contribute (or not) to behavioral change. In this way, these ''forces'' contribute to the extent that teachers learn and develop as a result of an intervention, and to all the processes that the intervention brings about during its deployment and afterwards. The individual teacher and his or her classroom behavior can be seen as such a complex dynamic system in itself.

A complex dynamic system is a system consisting of a number[1] of interacting elements or components that change each other's properties and consequently show self-organization, the emergence of self-sustaining states and patterns of states (Goldspink, 2007; Jörg, 2011; Van Geert & Steenbeek, 2010). This means that various influences occur that become coordinated and that explain the changes and stabilities in the individual teacher's learning process over time (Steenbeek & Van Geert, 2013). These influences are very diverse, and range from the teacher's own previous experiences with activities aimed at professional development to a teacher's work load, the external pressure a teacher experiences in the form of inspection reports and standardized school performance tests, and the importance a teacher assigns to the content to be learned. In short, the teachers' professional learning emerges out of the teacher-context dynamics.

---

[1] The number of interacting elements needed to constitute a complex system typically depends on the nature of the interactions. For instance, a system consisting of no more than three celestial bodies exerting gravitational forces onto each other already shows complex behavior. In the educational sciences, complex systems typically consist of one to a few dozen interacting elements. Macro scale social phenomena such as economics involve the interactions of many millions of producers and consumers.

Most current research in developmental psychology recognizes the importance of studying human learning and development in context (Van Geert, 2009). However, this context is usually seen as an independent and distal variable, e.g. as in studying the statistical association between the kind of school and the academic performance of children (Van Geert & Steenbeek, 2005). Using a complexity approach perspective, it is important to examine the context as a proximal, dynamic factor, meaning that the context is continuously created by the actions of the participants, and that the actions of the participants are continuously dependent on the created context. According to the complexity approach, context and person influence each other, as there is a "reciprocal determinism" between the two (Steenbeek & Van Geert, 2007). In this case, the school context in which the teacher functions must not be considered as a stable context, but as dynamic one, which changes itself over time in association with the changes in the teacher(s) and other components of the student in the school system.

An intervention or activity aimed at improving the teacher's behavior can be regarded as an important component that changes the context in which the teacher functions, and acts as a kind of "perturbation" of the complex dynamic system of school-student-teacher relationships. Note perturbing a complex system does not necessarily imply that the system will change. In fact, complex dynamic systems tend to retain their stability by counteracting such perturbations in order to damp out its effects. Whether or not a perturbation will indeed lead to a change in a complex system depends on the properties of the perturbation and the conditions under which it happens. Given the mutuality of person and context, it is clear that the exact way in which a particular intervention (perturbation) unfolds is not independent of the reactions of the teacher to the intervention, or of the teachers' input in the intervention. The teacher's 'simplex system', i.e., the teacher's connected set of beliefs, representations, values, emotions, habits, practices and material tools plays a central role in this process. This 'simplex system' serves as a simplifying representation of the overarching complex system in which a person participates, and organizes the participants' actions (for more information about the notion of 'simplex systems', we refer to Van Geert & Steenbeek (2014)).

That is to say, even if we guard the faithful implementation of a particular intervention, its concrete form over time will unavoidably be influenced by the teacher(s) and by the effect the intervention has on the teacher(s) and other significant parts of the school context. At the same time, the intervention changes the teacher as a learning individual as well. In this process of change, it is important to examine the *fit* between a specific intervention with a specific teacher in a specific context with a specific provider of the intervention. This fit is not only an idiosyncratic process, which means that the outcome is specific for this specific individual (or specific team of teachers), but also a dynamic process, in that it changes as a result of the interactions.

With regard to the effect of interventions for teachers in a specific school context, this means that the (above described) intervention-specific aspects, teacher-specific aspects, context-specific aspects and aspects regarding implementation interact and combine in various ways, in such a way that a particular intervention can have multiple outcomes, dependent on the dynamics of the specific context-teacher relationship. The fact that average results can be seen for a particular sample or population of individuals does not imply that this average result is true for each member of the sample or population (Jarvis & Pell, 2004; Rose, Rouhani, & Fischer,

2013). In order to understand how and to what extent interventions work, they should be studied in the context of actual intervention processes taking place in individual teachers (Barlow & Nock, 2009; Molenaar & Campbell, 2009). In fact, what counts is the way in which the teacher assimilates something that is in principle a perturbation of the ongoing daily process, namely the intervention. The assimilation can take several dynamic forms. For instance, it may be the starting point for a positive change cycle. It may also cause resistance and changes that actually compensate for the attempt to change the current attractor state, etc. Thus, from a complexity viewpoint, implementation is the effect that the introduction of the perturbation (in this case the intervention) has on the current state of the 'teacher's system'.

As can be seen in figure 2, a reciprocal influence exists between a teacher and every single group of aspects that plays a role in the intervention. This figure illustrates how a teacher and all aspects influence each other during the intervention. The arrow shows how the teacher and all aspects influence each other. This figure stresses the importance of the fit between a teacher and each single group of aspects in order for an intervention to be (more or less) effective for a particular teacher. Secondly, the figure stresses the fact that the influences exist between the teacher and all the different aspects at the same time, and that the influences are not one-directional, but multi-directional.

Thirdly, Figure 0 shows that, when describing a concrete, locally situated intervention in a complex dynamic system, it is difficult to explain the effective aspects in isolation. In a complex dynamic system, these groups of aspects, for instance, context based and teacher specific aspects are all intertwined during the actual process of an intervention, and the effectiveness for an individual teacher dynamically emerges from this interaction.

This implicates that, when conducting an intervention study, researchers should take into account that the separate aspects are important, but also that all these aspects are intertwined in a real world situation. This means that the combination of both approaches is meaningful. On the one hand, the standard sample-based average effect of the intervention with various aspects as largely independent variables contributing linearly to the statistical group effect can inform us about the net effect of an intervention for a particular group or population, and is typically important for the level of policymaking. On the other hand, the case-based dynamic understanding of how the actual intervention process leads to a particular result for a particular person (or for a particular school or team of teachers) is typically important for the level of actual intervening, i.e., this approach provides information about how actual interventions can be carried out and adapted within the limits of implementation integrity. In the third main section of this paper, we shall now illustrate the importance of both approaches by describing a multiple case study, in which two primary school teachers participated in an intervention aimed at teachers' professional development, especially with regard to improving the quality of their science lessons.
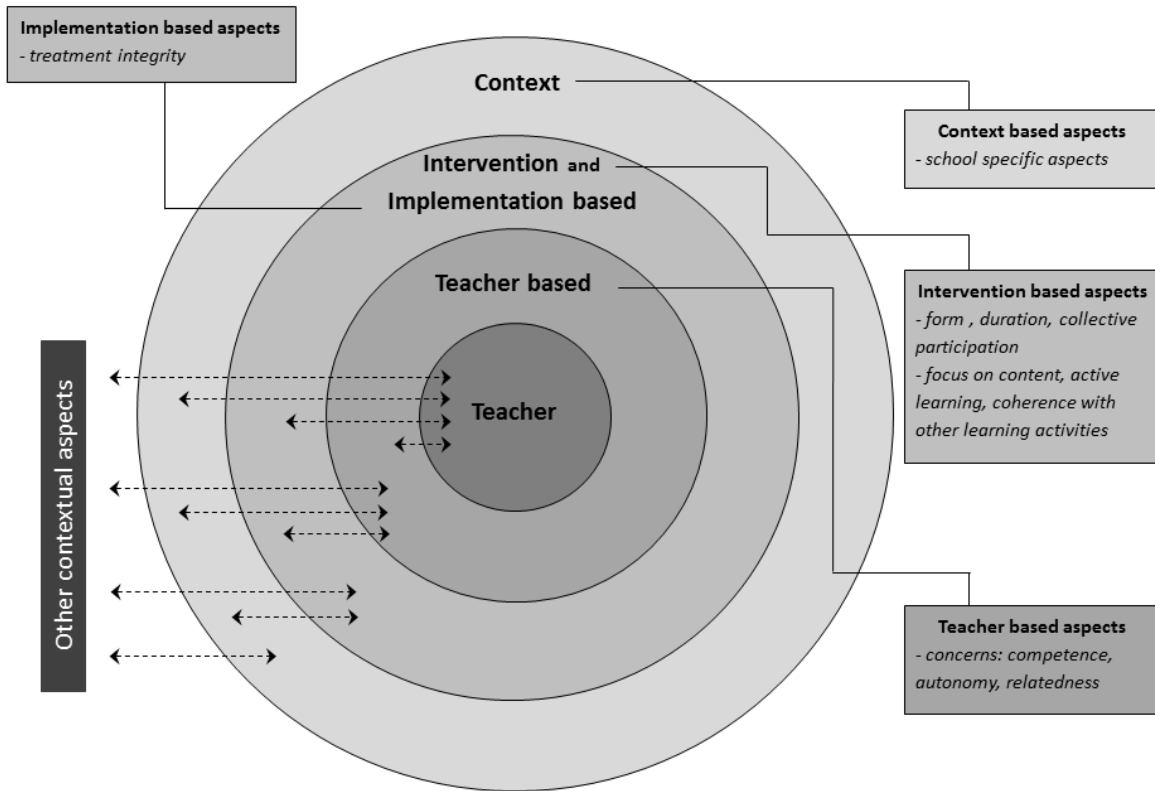
Figure 0. An intervention using a complexity approach

3 Empirical Illustration: The Design of a Science Teaching Intervention and its Effects, Using a Complexity Approach

3.1 Theoretical background

Various studies have shown that an important determinant for young students' success in science education is the students' curiosity, a quality that young children start with naturally (Jirout & Klahr, 2012; Loewenstein, 1994). However, once young children go to school, this curiosity seems to vanish. Engel (2009) observed that teachers hardly encourage (and sometimes even discourage (Martin-Hansen, (2009)) inquiry and curiosity in the classroom. This lack of stimulation might explain the diminishing interest in learning science in the first school years of children (Osborne, Simon, & Collins, 2003). This finding implies that it is important to start from the early school years on with encouraging children's natural curiosity and scientific reasoning, and thus, with acquiring positive science experiences. Teachers play a significant role in providing these positive experiences, and therefore it is important to design and implement professional development trajectories aimed at improving teacher's knowledge, skills, and attitudes in teaching science.

In our study, the goal was to design an intervention for improving the quality of teachers' science teaching, based on principles from (School) Video Interaction Guidance (Kennedy, Landor, & Todd, 2011**; Van den Heijkant et al., 2004). The aim of the intervention was to enable lower grade teachers, without preliminary experience in teaching science, to teach science in their own classroom by carrying out small scientific experiments. Our hypothesis was that in the concrete situation of each particular teacher, the effective aspects would work together in a specific way, and in fact amplify each other to produce a specific effect of the intervention. We also assumed that these processes of self-amplification or self-reduction are likely to be highly idiosyncratic, and thus will show high inter-individual variability. We expected that there would be an effect of the intervention in each particular teacher, but that this effect would greatly differ between the teachers due to the idiographic process, i.e., specific interplay of the intertwining aspects within each teacher.

3.2 *The content of the intervention*

When developing the intervention, we started from the following basic assumptions. The first assumption is that children are curious about their environment. The second assumption is that teachers have some knowledge of basic science principles, for instance, of floating and sinking and gravity[2]. This implies that it should be possible for teachers to start with simple experiments in the classroom that do not require specific teacher training. The emphasis of the intervention was not on scientific content knowledge, but on pedagogical content knowledge in science lessons (Smith, 1999). Teachers needed to learn and experience how to enhance children's curiosity and scientific reasoning during these lessons. Therefore, teachers received as part of the intervention information they were not familiar with beforehand about questioning (Oliveira, 2010), about the empirical cycle, and about scaffolding., in a two hour theory session. Scaffolding (Granott et al., 2002; Mattanah, Pratt, Cowan, & Cowan, 2005; Van Geert & Steenbeek, 2005) was explained as a way of helping pupils to reach a higher level than they could reach on their own by being given parts of the solution to a problem. The aim of scaffolding is that pupils ultimately can give solutions independent of a teacher's help. The empirical cycle (De Groot, 1961) was explained as a five step inquiry scheme to provide structure to a science lesson. The empirical cycle, as applied to what young pupils can do in a science lesson, consists of the following steps: first asking your research question (for example why does a paperclip sink whereas a big piece of wood floats) second hypothesizing (for example, we think that the paperclip sinks because it is made out of iron), third designing a research model (for example, we take several iron objects, including a small metal box, and we predict that they will all sink), fourth observing (for example, try each metal object separately and observe whether it sinks are not), explaining the observed result (for example, because some metal objects do not like water, they will float). The cycle can then be repeated by taking the last step as its starting point, for example, the pupils ask how they could know whether

---

[2] It should be noted that teachers, and many other educated adults for that matter, can hold serious misconceptions about physical science concepts, such as those involving Archimedes law. However the basic science principles involved in the current science lessons were truly basic and intuitive, and none of the teachers showed any misconceptions on this level.

certain metal objects like or dislike water. In the current example, the explanation is not correct from a scientific point of view, and repeating the empirical cycle, supported by questions asked by the teacher, should help children to move towards better explanations and better predictions

After the theory lesson, teachers prepared a science lesson four times to be given in their own classroom with a small group of children. Each lesson was recorded on video. Afterwards the coach selected relevant video fragments with a focus on teacher questioning, the use of the empirical cycle, and scaffolding in the interaction with the children. These fragments were reflected upon by teacher and coach together using the principles of Video Enhanced Reflective Practice (VERP) (Kennedy et al., 2011). The focus of the coaching sessions was to enable the teacher to discover in interaction with the coach, how to change behavior by discussing fragments that were representative of the teacher's behavior.

3.3 *The design of the intervention*

Important aspects of interventions that contribute to the effectiveness of an intervention and the way these aspects can best be addressed were specified above. The next sections describe how these aspects were incorporated in our science teaching intervention.

3.3.1 Intervention specific aspects

The form of the intervention was designed as training within the classroom with elements that have proven to be effective, as there is a direct connection between the teaching practice (Loucks-Horsley et al., 2010) and the use of direct performance feedback (Noell et al., 1997). The duration of the intervention was defined by the total period of time in which the intervention sessions were given and the number of contact hours. The intervention lasted five to eight weeks. Every coaching session consisted of a science lesson of 20 to 30 minutes in the teacher's own classroom, followed by 30-45 minutes of coaching. This duration was defined by a combination of time and cost effectiveness. Because of the importance of collective participation, the aim was to involve most teachers of the lower grades to participate. In the actual intervention, we had three participants in one school, and four participants in the other school, which gave the teachers the possibility to share their experiences.

The focus in the intervention was on improving teachers' pedagogical content knowledge, by means of active learning. That is, during the four lessons with subsequent coaching, teachers were given the possibility to put the learned framework into practice and to reflect on their own lesson with the help of a trained coach as described in 3.2. Moreover, the focus was on coherence with other learning activities. This aspect was not incorporated in the design of the intervention, although schools usually volunteered to participate because they viewed science as an important topic in the school curriculum. Thus, a broader interest in the topic implicated that more activities regarding science lessons are already being undertaken at the school, although not necessarily in coherence with other learning activities.

3.3.2 Teacher specific aspects

The teacher specific aspects were addressed by taking teachers' concerns, (autonomy, relatedness and competence) into account, so that teachers could feel more intrinsically motivated to use the newly learned content.

Firstly, the teachers' competence concern was addressed by presenting the teachers with an interactive lecture on the knowledge they need when teaching science. The aim was to let teachers experience that they already possessed enough content knowledge to teach science to young children, e.g. about floating and sinking, gravity and air pressure experiments. Furthermore, during the video coaching, the coach addressed not only learning points, but also showed the teachers video shots in which the teacher showed very competent behavior.

The autonomy concern was addressed by letting teachers voluntarily participate in the intervention. Schools were asked to include only teachers who were enthusiastic about teaching science and who wanted to learn how to improve their teaching skills at this topic. Furthermore, participating teachers were asked to teach science lessons of their own choice in their classroom during the trajectory. Moreover, teachers were asked to formulate their own learning goal, as autonomous goal setting allows them to be responsible for their own learning trajectory and leads to better results in the process of behavioral change (Koestner, 2008).

As for the relatedness concern, we asked schools to let more than one teacher of one school team participate, so that these teachers could experience and discuss the theory session together. In addition, the improved questioning and reacting to pupils contributed to the relationship between teacher and pupil.

3.3.3 Context specific aspects

Only school specific factors were included in the design of the intervention because they could be addressed directly. Organizational support in training, for instance, extra time for study and practice, and supervision and support by the school principal during the intervention's implementation and afterwards, were important activities within the design. School principals were asked to endorse the starting points of the intervention, agree with the time the intervention would take, and to support their teachers by providing substitute teachers in the class during the coaching sessions. The principal and the colleagues who did not participate in the intervention also had the opportunity to attend the theory sessions, so that all teaching staff had the same level of knowledge, even though not everyone participated in the overall coaching program. In school A, three teachers participated in the coaching program, and in school B four did, which made the sharing of mutual experiences possible.

3.3.4 Aspects regarding implementation

The intervention as a whole was clearly documented. A manual of the intervention was made available to enable dissemination of the intervention to other trainers and to ensure treatment integrity (Wetzels, Steenbeek, & Fraiquin, 2011).

4 Studying the Effect of the Intervention Using a Complexity Approach

The following multiple case study is part of the effect study that was conducted in order to examine the effectiveness of Video Feedback Coaching for Teachers (VFC-T), the coaching program for teachers supporting the teaching of science in the lower grades. The aim of this case study was to illustrate how, in the concrete situation of a particular teacher, the effective aspects work together in a specific way, and in fact either amplify each other or reduce the

influence of the intertwined aspects to produce a specific effect of the intervention. We also assume that these processes of self-amplification or self-reduction are likely to be highly idiosyncratic, and thus will cause high inter-individual variability between teachers.

4.1 Participants

Two teachers participated: Laura (age 46, teaching experience 23 years) from school A and Anne (age 53, teaching experience 22 years) from school B. Both teachers were representative of their co-workers in the same school as classroom teachers of pre-schoolers (age 4-6), and had almost the same knowledge of science and science education as well as experience in teaching science. The selection was based on information we collected with regard to age, knowledge of science teaching and experience in teaching science before the intervention. For the coaching trajectory they each chose four children from their class, children they considered representative for the whole class with regard to their age, performance level, and ethnic background. The pupils' age ranged from four to six years, and none of them were from an ethnic or cultural minority because in the part of the country where the study was carried out the number of children of ethnic or cultural minorities is relatively small.

4.2 Data collection, variables and analyses

The study uses a mixed method, combining quantitative and qualitative analyses (Creswell & Plano Clark, 2011). Data were collected by using video recordings of the classroom activities, which were captured with two digital camcorders, one focusing on the teacher, and the other one on the participating pupils. Afterwards, the video recordings were coded, using a coding scheme for the variables "amount of questions" and "level of reasoning".

Two coding schemes were developed for coding teachers' (Wetzels, Steenbeek, & Van Geert, 2015) and pupils' utterances (Wetzels, Steenbeek, & van Geert, 2015). The coding scheme regarding teachers' utterances was operationalized as the variable 'teachers' scientific reasoning eliciting questioning' (TSEQ), and it was developed using literature that stresses the importance of questions regarding the empirical cycle for science learning (De Groot, 1961; Dejonckheere, Van De Keere, & Mestdagh, 2009; Engel & Randall, 2008; Oliveira, 2010). All questions were selected and coded, if possible, in terms of the steps of the empirical cycle (see section 3.2): 'knowledge questions', 'prediction questions', research design questions', 'observational questions' and 'questions about the explanation'. Additional categories of questions were: 'follow up questions', 'other questions relating to the content', and 'other questions'.

The coding scheme (table 1) regarding pupils' utterances was operationalized as the variable 'pupils' level of reasoning' (from now on PLS), and is based on skill theory (Fischer & Bidell, 2006). It describes the cognitive level in pupils' utterances on an interval scale, according to their skill level. For each lesson, which lasted about 20-30 minutes, the highest level one of the pupils reached was coded per minute, and subsequently the mean score for each lesson was calculated. When children said nothing during a minute, it was coded as zero because in this case the teacher did not elicit children's reasoning at all. At the least complex skill level we coded a 1. A 2- 5 code indicated increasing skills. (Meindertsma, Van Dijk, Steenbeek, & Van Geert, 2012; Rappolt-schlichtmann, Tenenbaum, Koepke, & Fischer, 2007)

Table 1: coding scheme for pupils' level of scientific reasoning (PLS)

Code	Skill level	**Description**	**Example**
0		No answer is elicited by the teacher	
1	Sensorimotor mapping	A child observes characteristics of an object	This is a blue pencil
2	Sensorimotor systems	Child states a relationship between action and result	The pencil floats because you placed it in the water
3	Single representation	Child refers to one part of the explaining mechanism	The pencil floats because it is small
4	Representational mapping	Child refers to two or more parts of the explaining mechanism	The pencil floats because it is small and light
5	Representational system	Child refers to all explaining mechanisms	The pencil floats because it is light for its size in the water

Differences between pre and post-intervention and pre-intervention and intervention lessons for each individual case were analysed using Monte Carlo permutation analyses (Todman & Dugard, 2001), a non-parametric permutation test, particularly suitable for small sample sizes. The probability that results are caused by chance alone was estimated by simulating the chance-alone condition. This simulation is done by randomly shuffling all results numerous times (1000 or more). The original result is then compared with the distribution of the results obtained by the random model, using a p-value of .05, to determine the probability of whether the observed result for each particular teacher might be due to chance alone.

After the last post-intervention, the participants were interviewed by using a semi-structured interview technique. Questions focused on the way the teachers had experienced the coaching trajectory and on the changes the teachers observed regarding their own behavior and regarding the children in the classroom.

4.3 Results

4.3.1 Quantitative findings – TSEQ and PLS

The results of Laura (as represented in Figure 1a) show that the amount of TSEQ increased significantly after the two pre-intervention lessons and the theory session. This higher amount of TSEQ stayed at a higher level during the four coaching sessions (p=0.02). However, it dropped sharply two months after the coaching sessions, during the post-intervention lessons, which implies that no statistically significant results were seen in the longer term (p=0.50). With regard to the variable PLS (Figure 1b), a significant increase was found during the coaching

lessons (*p*=0.02) in comparison with the two pre-intervention lessons, and this higher level was continued during the post-intervention lessons (*p*=0.03).

On the contrary, with Anne (Figures 2a and 2b), no significant changes occurred during the four coaching sessions, nor two months after the coaching sessions, during the post-intervention lesson, both regarding the amount of TSEQ (short term *p*=0.87, long term *p*=0.69) and the PLS (short term *p*=0.89, long term *p*=0.77).

Laura

1a. Amount of emprical cycle questions

1b. Mean level of reasoning pupils

Anne

2a. Amount of emprical cycle questions
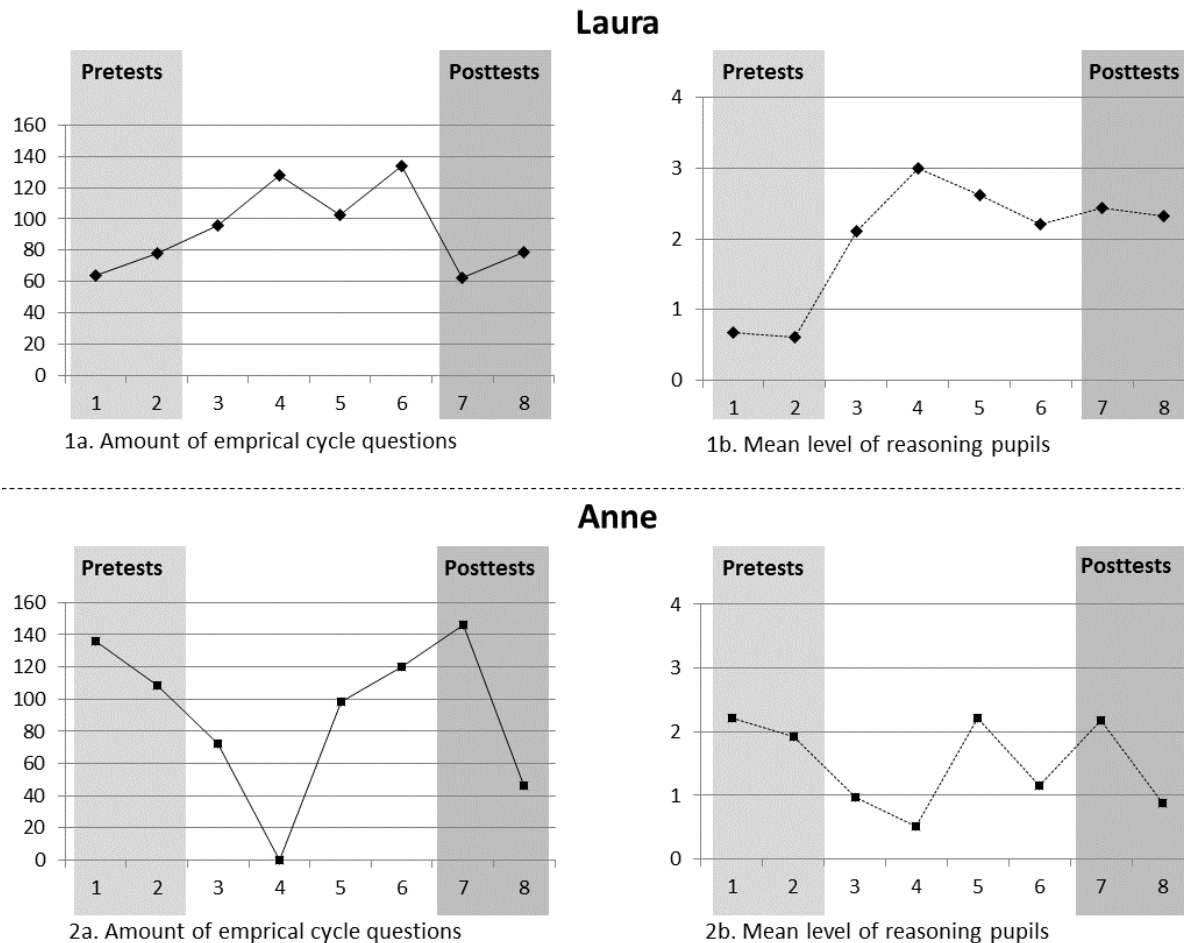
2b. Mean level of reasoning pupils

Figure 1a, 1b, 2a and 2b. Results Laura and Anne

In the interview, both teachers were positive about the form of the intervention in that they gave the intervention a very positive rating (Anne gave a 7, Laura gave a 8 on a scale of 1-10). However, the quantitative variables show a high amount of inter-individual variability with regard to the teachers' questions, and the pupils' 'level of reasoning'. Presumably related to the differences between the teachers, Anne shows a higher amount of intra-individual variability than Laura.

4.3.2 Comparison of the two trajectories regarding the influencing aspects and the intertwining between them

In the following we qualitatively describe (by giving examples) how all aspects have an influence on the effectiveness of the intervention in the multiple case study, not only alongside

each other, but also intertwiningly influencing each other, using the interview data of both teachers.

The science teaching intervention was designed with several effective aspects incorporated (*cf* the design of intervention section above). The intervention-specific aspects, both structural and substantive, were the same for both teachers, as well as the treatment integrity, as the coach was the same for both teachers and used the same documentation.

Some examples of differences between the two trajectories can be clearly observed: for instance, the difference regarding the fit between school and intervention. Laura recognized that the intervention fit extremely well with the cornerstones of her school's teaching, namely the Reggio Emilia approach (Edwards, Gandini, & Forman, 1993). One of the aspects in the coaching program was scaffolding, requiring a fit with the level of the child and providing specific attention to each individual child. This manner of working with specific attention for individual children is a key feature of the Reggio approach, and Laura was happy to see how both merged together. That is to say, important content aspects of the school background and of the intervention were clearly consistent with one another, which might have resulted in amplifying Laura's enthusiasm. On the contrary, in Anne's school, science was an important subject; however, most teachers in that school were not required to participate in the science program themselves, since the responsibility was given to a specially appointed science teacher, which likely led to self-reduction. That is, Anne felt like the coaching did not help her with the regular teaching, probably did not see any result that she considered valuable, and most likely, for this reason, concluded that the coaching required too much of her time, as she repeatedly said during the interview.

Regarding the teacher specific aspects, clear differences between the teachers could be seen with regard to feelings of competence and autonomy. During the pre-intervention lessons, Laura did not feel competent in teaching science, but she wanted to gain more knowledge and skills on this subject. She was very critical at the start of the theory session, but she became increasingly enthusiastic, and she understood the advantages of teaching science in this challenging way. On the contrary, Anne had some experience with science education, but primarily with the science coordinator giving science lessons in her classroom. During the theory lesson she mentioned that she wanted to help with the research, but that she felt she already knew enough about science teaching, although she had never put this knowledge in practice herself. This indicated a clear difference in the fit between the intervention and the teacher's simplex systems (see: Van Geert & Steenbeek, 2014).

Regarding their feeling of autonomy, it was clear that Laura was positive about the opportunity to participate: "Our principal told us about the opportunity. And the topic seemed very interesting to me and my colleagues, so we said yes". In contrast, Anne's answer was: "Oh, we had to do this. All the other colleagues said no, and someone had to participate. And then I said yes, but it took much more time than I expected". In fact, it seems that Laura's decision was based on the content of the program, and enhanced her feeling of autonomy, whereas Anne decision was not content-based, and even seemed to threaten her feeling of autonomy.

Furthermore, both teachers set their own learning goals. During the interview, Laura said: "the most important learning goal for me was to hold myself away from talking and answering, and give the children room for their answers. And I do much better at this now."

She could still recall her learning goal two months later, and could describe what she had learned. Anne could not recall what her goal had been. In her own word: "it was such a long time ago". Stated in terms of concerns, Laura seems to address her basic concerns through the coaching. On the contrary, Anne very clearly explained that the coaching did not address any of the concerns she finds interesting, in the sense that she complained about the amount of time it required, without identifying any feeling of positive results (the concern that she expressed was an external concern, namely the willingness to participate in scientific research, not the concern of changing her own competence) (Steenbeek & Van Geert, 2013).

Regarding the school-specific aspects, two aspects show clear differences between Laura and Anne, namely the nature and extent of the school support and the school wide approach. In Laura's school, the principal was involved in the process of teaching science. The principal wanted the school to participate in the intervention research because she intended to start with a few teachers in the lower grades, and to eventually extend this to the rest of the teachers in the school. To express her commitment and make this observable within the organization, she participated in the theory session and in the discussion on teaching science that emerged during that session. She ensured that teachers would have time to participate and that their pupils would be taken care of during the intervention time. Laura and her co-workers felt supported and they understood how important the intervention was for the future of the school. In contrast, in Anne's school, the teachers participated in the research project because the science coordinator asked them to. The school principal did not know about it, and during the theory session the teachers participated alone, without the science coordinator or the principal. During the coaching sessions, Anne's replacement in the classroom was taken care of, however Anne felt the trajectory took too much of their time.

Regarding the school-wide approach, the general idea at both schools was that in the end all teachers would learn how to teach science. In Laura's school, all four lower grade teachers participated, because the principal had asked them to be an example for the upper grade teachers, who would follow the trajectory at a later time. The principal had asked them in advance to think of a way to communicate the learned content to their upper grade colleagues at the end of the coaching program in order to inspire them too for the science teaching program. In Anne's school on the contrary, only three of six lower grade teachers participated, and no plans were made regarding whole school dissemination.

In the previous section, the intertwining of different aspects shows that in the case of Laura, she was enthusiastic about the intervention because its goals fit with the school's Reggio Emilia approach, the principal was enthusiastic, as well as her co-workers. Her personal concerns, school specific aspects and intervention specific aspects worked intertwiningly together in a positive spiral to produce a relatively positive result. However, in the case of Anne, the intertwining of aspects unravelled very differently, i.e., she was not enthusiastic nor was her principal, which led to a relatively negative result.

5. Discussion

The aim of this article was to give a theoretical overview of effective aspects of interventions regarding teachers' professionalizing, to re-interpret the effectiveness of these aspects from a complexity view, and to illustrate this complexity view with a multiple case study.

In principle, the results presented in our case studies could be interpreted in terms of simple, linear additions of independent variables, such as variability in teacher ability, in the student's intelligence, variation in motivation, variation in students' experience and so forth. The complexity view, on the other hand, provides an interpretive framework focusing on the reciprocal influences of all these, and other factors and components. In these two case studies, we have illustrated how intervention specific aspects, teacher specific aspects, context specific aspects and implementation specific aspects intertwine, and act as a complex dynamic system, that is to say as a system of interacting components, leading to self-organizing patterns of asking questions, giving explanations, involvement in the task, the teachers willingness and effort to improve his or her ability to teach science, and so forth. From a complexity view, we focus on an interaction-dominant analysis (Van Orden, Holden, & Turvey, 2003) of intervention effects , which entails the analysis of how these adjacent processes interact and change each other's dynamics, and in which each interaction pays attention to the specific conditions of the adjacent processes. In this study, this means that we first looked at how the intervention works at the level of real-time activities of individual teachers in the context of their schools, and subsequently we looked at the dynamic interplay between variables that matter at this level of analysis, such as individual concerns and individual evaluations of the value of the obtained effects. In this way, an intervention can be seen as a process, a sequence of activities of a teacher, her children, and an intervener. How plausible is it that such a sequence of activities in which the teacher, children and intervener react to each other in a variety of ways can be explained by a superposition of influences exerted by many independently operating variables? In our view, such an explanation is highly implausible: it does not allow us to interpret the observable interactions that take place during these processes. The alternative theory – and most likely the only alternative present —for interpreting processes of change in teacher and pupil activities intertwined is a theory that treats such processes as examples of interactions between components, as conceptualized in the general theory of complex dynamic systems.

In large scale research it is impossible to pay attention to all the influences that can play a role in an intervention process. From a complexity perspective the description of all aspects is not possible because of the multitude of possible aspects, and because these aspects continuously change over time. It is important, however, to recognize that all contextual aspects influence each other intertwiningly and form positive or negative feedback loops. These effects need to be taken into account when assessing the effectiveness of an intervention, or when providing an intervention. Moreover, when designing an intervention from a complexity perspective, attention must be paid to the fit between intervention, teacher, provider, and the context in the broadest sense. The notion of the intertwining and reciprocal relations between all aspects requires flexibility and an open mind of the trainer. The trainer himself is in a complex dynamic system, embedded in the process itself, and therefore does not function as an external factor. While maintaining the intervention integrity, the trainer must be able to make adaptations at the right time and at the right place, by focusing on the intervention as an idiosyncratic, individual-based process. We speculate that the example at the end of the

previous section could have worked out very differently if the principal had been very positive about the intervention. If she or he had shown this enthusiasm by participating in the theory session, Anne's concerns could have changed her motivation to more intrinsic motivation for participating, instead of the extrinsic motivation she had. This change into intrinsic motivation could also have had its influence on her idea about the learning goals and the length of the intervention. The results and the effectiveness of the intervention could have displayed a more positive picture for Anne in this situation.

Some limitations of this research need to be mentioned. Firstly, only a part of the influencing context is discussed. More aspects are important when studying an intervention's effectiveness (Boelhouwer, 2013; Boomstra, 2014), and these aspects can vary from the enthusiasm of the children, to bureaucracy that enforces teachers to unwanted work, and to the influence of a teacher's sick child at home. An exhaustive description, however, of the influencing aspects is not necessary in a complex dynamic system because what matters are the intertwining processes and the interactions between the influencing aspects. Secondly, the empirical illustration is limited to two cases, in which not all possible aspects are visible, and therefore cannot be taken into account. However, the results of this study show the importance of a complexity approach when intervening.

Designing and implementing a sustainable intervention for professional development is not an easy task, not least because most systems resist change in the long run. This kind of resistance can only be understood from a complexity viewpoint concerning the role of interventions, in that all aspects are seen as components that sustain each other (Frenken, 2006). It can take up to three to five years for teachers to fully implement a new practice or program, and therefore expecting change in a short period of time is unrealistic (Loucks-Horsley et al., 2010). This can be regarded as a plea for long-term intervention processes, as well as longitudinal studies in order to examine the effects of interventions over a longer timeframe.

The complexity approach to interventions has some important implications for future empirical research, as well as for the design of interventions. In future research, results of small scale and large scale research should be combined, in order to obtain a better insight into all relevant aspects and their effect on teachers' behavior. In addition, research should concentrate on the effects that all contextual aspects have intertwiningly. Once more knowledge is available concerning how aspects influence each other contextually, this knowledge can be used to design more effective interventions and to implement them more effectively with higher chances that the system incorporates the new influence and that effects will last.

References

Adelman, H. S., & Taylor, L. (2003). On sustainability of project innovations as systemic change. *Journal of Educational and Psychological Consultation*, *14*, 1–25. doi:10.1207/S1532768XJEPC1401

Armour, K. M., & Yelling, M. R. (2004). Continuing professional development for experienced physical education teachers: Towards effective provision. *Sport, Education and Society*, *9*, 95–114. doi:10.1080/1357332042000175836

Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science*, *4*, 19–21. doi:10.1111/j.1745-6924.2009.01088.x

Barth, R. S. (2004). *Learning by heart*. San Francisco: Jossey-Bass.

Berman, P., McLaughlin, M. W. A., Bass, G., Pauly, E., & Zallman, G. (1977). *Federal programs supporting educational change, vol VII: Factors affecting implementation and continuation*. Santa Monica, CA: The Rand Corporation. Retrieved from http://www.osti.gov/servlets/purl/1025774

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

Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership*, *57*, 28–33.

Boelhouwer, M. D. (2013). *Tussen weerstand en weerbaarheid en andere recepten*. University of Groningen.

Boomstra, N. (2014). *Read all about it !* University of Groningen.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12. doi:10.3102/0013189X032006003

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, *102*, 294–343. doi:10.1111/0161-4681.00057

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Day, C. (1999). *Developing teachers. The challenge of lifelong learning*. New York: Taylor & Francis Inc.

De Groot, A. D. (1961). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen*. 's Gravenhage: Mouton & Co.

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. (E. L. Deci & R. M. Ryan, Eds.). Rochester (NY): The University of Rochester Press.

Dejonckheere, P. J. N., Van De Keere, K., & Mestdagh, N. (2009). Training the scientific thinking circle in pre- and primary school children. *The Journal of Educational Research*, *103*, 1–16. doi:10.1080/00220670903228595

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*, 181–199. doi:10.3102/0013189X08331140

Dewey, J. (1938). *Experience and education*. New York: Collier Books.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542–575.

Dinham, S., & Scott, C. (2004). International patterns of teacher satisfaction and motivation in Australia, England, New Zealand, United States of America, Cyprus and Malta: The role of context and the "third domain." In *Paper presented at the British Educational Research Association Annual Conference, University of Manchester, 16-18 September 2004*.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–50. doi:10.1007/s10464-008-9165-0

Edwards, C., Gandini, L., & Forman, G. (Eds.). (1993). *The hundred languages of children: The Reggio Emilia approach to early childhood education*. Norwood, NJ: Ablex.

Engel, S. (2009). Is curiosity vanishing? *Journal of American Academy of Child and Adolescent Psychiatry*, *48*, 777–779. doi:10.1097/CHI.0b013e3181aa03b0

Engel, S., & Randall, K. (2008). How teachers respond to children's inquiry. *American Educational Research Journal*, *46*, 183–202. doi:10.3102/0002831208323274

Evans, L. (2002). What is teacher development? *Oxford Review of Education*, *28*, 123–137. doi:10.1080/03054980120113670

Fabiano, G. A., Vujnovic, R. K., Waschbusch, D. A., Yu, J., Mashtare, T., Pariseau, M. E., … Smalls, K. J. (2013). A comparison of workshop training versus intensive, experiential training for improving behavior support skills in early educators. *Early Childhood Research Quarterly*, *28*, 450–460. doi:10.1016/j.ecresq.2012.12.001

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, *28*, 465–498.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. In R. M. Lerner & W. Damon (Eds.), *Theoretical models of human development. Handbook of child psychology* (pp. 313–399). New York: Wiley.

Frenken, K. (2006). *Innovation, evolution and complexity theory*. Cheltenham (UK): Edward Elgar Publishing Limited.

Frijda, N. H. (1986). *The emotions*. Cambridge University Press. Retrieved from http://www.oxfordscholarship.com/oso/public/content/philosophy/0199253048/toc.html

Fullan, M. (2007). *The new meaning of educational change*. New York, NY US: Teachers College Press.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journaleducational Research Journal*, *38*, 915. doi:10.3102/00028312038004915

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., … Hurlburt, S. (2011). *Middle School Mathematics Professional Development Impact Study*.

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review*, *31*, 79–88. doi:10.1016/j.cpr.2010.09.007

Gersten, R., Taylor, M. J., Keys, T. D., Rolfhus, E., & Newman-Gonchar, R. (2014). Summary of research on the effectiveness of math professional development approaches. (REL 2014–010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Goldenberg, C., & Gallimore, R. (1991). Changing teaching takes more than a one-shot workshop. *Educational Leadership*, *49*, 69–72.

Goldspink, C. (2007). Transforming education: Evidential support for a complex systems approach. *E:CO*, *9*, 77–92.

Goodrum, D., Hackling, M., & Rennie, L. (2001). *The status and quality of teaching and learning of science in Australian schools*. *Department of Education, Training and Youth Affairs*. Canberra.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, *39*, 3–35.

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). Cambridge: Cambridge University Press.

Grigg, J., Kelly, K. A., Gamoran, A., & Borman, G. D. (2012). Effects of two scientific inquiry professional development interventions on teaching practice. *Educational Evaluation and Policy Analysis*, *35*, 38–56. doi:10.3102/0162373712461851

Guskey, T. R. (2003). What makes professional development effective ? *Phi Delta Kappan*, *84*, 748–750. doi:10.3102/00028312038004915

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*, *33*, 665–79. doi:10.1007/s10802-005-7646-2

Harwell, S. H. (2003). *Teacher professional development: It's not an event, it's a process*. Waco, Texas: CORD.

Hattie, J. (2003). *Teachers make a difference: What is the research evidence ? Australian Council for Educational Research Annual Conference on Building Teacher Quality*.

Hattie, J. A. C. (2009). *Visisble learning. A synthesis of over 800 meta-analyses relating to achievement*. New York, NY US: Routledge.

Hetherington, L. (2013). Complexity thinking and methodology: The potential of "complex case study" for educational research. *Complicity: An International Journal of Complexity and Education*, *10*, 71–85.

Imperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development*. Wellington, New Zealand.

Jarvis, T., & Pell, A. (2004). Primary teachers' changing attitudes and cognition during a two-year science in-service programme and their effect on pupils. *International Journal of Science Education*, *26*, 1787–1811. doi:10.1080/0950069042000243763

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, *32*, 125–160. doi:10.1016/j.dr.2012.04.002

Jörg, T. (2011). *New thinking in complexity for the social sciences and humanities: A generative, transdisciplinary approach*. New York, NY: Springer.

Keiny, S. (1994). Constructivism and teachers' professional development. *Teaching and Teacher Education*, *10*, 157–167.

Kennedy, H., Landor, M., & Todd, L. (Eds.). (2011). *Video Interaction Guidance: a relationship-based intervention to promote attunement, empathy and bellbeing*. London: Jessica Kingsley Publishers.

Koestner, R. (2008). Reaching one's personal goals: A motivational perspective focused on autonomy. *Canadian Psychology/Psychologie Canadienne*, *49*, 60–67. doi:10.1037/0708-5591.49.1.60

Lieberman, A. (Ed.). (1996). Practices that support teacher development: Transform- ing conceptions of professional learning. In M. W. McLaughlin & I. Oberman (Eds.), Teacher learning: Newpolicies, new practices (pp. 185-201). New York: Teachers College Press

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*, 75–98. doi:10.1037/0033-2909.116.1.75

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin.

Martin-Hansen, L. (2009). I*nquiry pedagogy and the preservice science teacher*. Amherst, NY: Cambria Press.

Mattanah, J., Pratt, M., Cowan, P., & Cowan, C. (2005). Authoritative parenting, parental scaffolding of long-division mathematics, and children's academic competence in fourth grade. *Journal of Applied Developmental Psychology*, *26*, 85–106. doi:10.1016/j.appdev.2004.10.007

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist*, *59*(1), 14–9. doi:10.1037/0003-066X.59.1.14

Meindertsma, H. B., Van Dijk, M. W. G., Steenbeek, H. W., & Van Geert, P. L. C. (2012). Application of Skill Theory to compare scientific reasoning of young children in different tasks. *Netherlands Journal of Psychology*, *67*, 9–19.

Minnaert, A., Boekaerts, M., & de Brabander, C. (2007). Autonomy, competence, and social relatedness in task interest within project-based education. *Psychological Reports*, *101*, 574–586. doi:10.2466/PR0.101.2.574.586

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112–117. doi:10.1111/j.1467-8721.2009.01619.x

Noell, G. H., & Gansle, K. A. (2009). Moving from good ideas in educational systems change to sustainable program implementation: Coming to terms with some of the realities. *Psychology in the Schools*, *46*, 78–89. doi:10.1002/pits.20355

Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does Treatment Integrity Matter ? A Preliminary Investigation of Instructional Implementation and Mathematics Performance. *Journal of Behavioral Education*, *11*(1), 51–67.

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly*, *12*, 77–88. doi:10.1037/h0088949

Oliveira, A. W. (2010). Improving teacher questioning in science inquiry discussions through professional development. *Journal of Research in Science Teaching*, *47*, 422–453. doi:10.1002/tea.20345

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*, 376–407. doi:10.3102/0034654311413609

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*, 1049–1079. doi:10.1080/0950069032000032199

Palmer, D. H. (2002). Factors contributing to attitude exchange amongst preservice elementary teachers. *Science Education*, *86*, 122–138. doi:10.1002/sce.10007

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clincal Psychology: Science and Practice*, 12, 365–383. doi:10.1093/clipsy/bpi045

Pressley, M., Graham, S., & Harris, K. (2006). The state of educational intervention research as viewed through the lens of literacy intervention. *The British Journal of Educational Psychology*, 76, 1–19. doi:10.1348/000709905X66035

Rappolt-Schlichtmann, G., Tenenbaum, H. R., Koepke, M. F., & Fischer, K. W. (2007). Transient and robust knowledge: Contextual support and the dynamics of children's reasoning about density. *Mind, Brain and Education*, 1, 98–108.

Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education*, 8, 5–35.

Rose, L. T., Rouhani, P., & Fischer, K. W. (2013). The science of the individual. *Mind, Brain, and Education*, 7, 152–158. doi:10.1111/mbe.12021

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78. doi:10.1037/0003-066x.55.1.68

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the new Reform. *Harvard Educational Review*, 57, 1–21.

Smith, D. C. (1999). Pedagogical content knowledge in elementary science. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge*. Dordrecht, The Netherlands.

Steenbeek, H. W., & Van Geert, P. L. C. (2007). A theory and dynamic model of dyadic interaction: concerns, appraisals, and contagiousness in a developmental context. *Developmental Review*, 27, 1–40. doi:10.1016/j.dr.2006.06.002

Steenbeek, H. W., & Van Geert, P. L. C. (2013). The emergence of learning-teaching trajectories in education: A complex dynamic systems approach. *Nonlinear Dynamics, Psychology, and Life Sciences*, 17, 233–67.

Todman, J. B., & Dugard, P. (2001). *Single-case and small-n experimental designs: a practical guide to randomization tests*. Mahwah (NJ): Erlbaum.

Van den Heijkant, C., Quak, G., Van Swet, J., Vloet, K., De Vos, M., & Van der Wegen, R. (2004). *School Video Interactie Begeleiding*. Antwerpen-Apeldoorn: Garant.

Van Geert, P. L. C. (2009). Nonlinear complex dynamical systems in developmental psychology. In S. J. Guastello, M. Koopmans, & D. Pincus (Eds.), *Chaos and Complexity in Psychology; the theory of nonlinear dynamical systems* (pp. 242–281). Cambridge: Cambridge University Press.

Van Geert, P. L. C., & Steenbeek, H. W. (2005). The dynamics of scaffolding. *New Ideas in Psychology*, 23, 115–128. doi:10.1016/j.newideapsych.2006.05.003

Van Geert, P. L. C., & Steenbeek, H. W. (2010). Networks as complex dynamic systems: Applications to clinical and developmental psychology and psychopathology. *Behavioral and Brain Sciences*, 33, 174–175.

Van Geert, P. L. C., & Steenbeek, H. W. (2014). The good, the bad and the ugly? The dynamic interplay between educational practice, policy and research. *Complicity: An International Journal of Complexity and Education*, 11, 22–39.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology*, 132(3), 331–50. doi:10.1037/0096-3445.132.3.331

Wetzels, A. F. M., Steenbeek, H. W., & Fraiquin, M. (2011). *TalentenKracht in de klas: Een coachingsprogramma voor leerkrachten van groep 1-4*.

Wetzels, A. F. M., Steenbeek, H. W., & van Geert, P. L. C. (2015). Primary science teaching: behavior of teachers and their pupils during and after a coaching program.

Wetzels, A. F. M., Steenbeek, H. W., & Van Geert, P. L. C. (2015). The effect of a coaching program on teachers' well-being in science and technology lessons in primary education.

About the Authors

Annemie Wetzels gained a master's degree in clinical and developmental psychology at the University of Groningen, where she defended her PhD thesis in 2015. For this PhD thesis she has developed a Video Feedback Coaching program for preschool and grade 1-2 teachers. This PhD-project is part of the Curious Minds program and aims at improving teachers' skills for stimulating children's talented science learning. By using a complexity approach she studies the effects of this coaching program on teachers and children. Currently, she works as an independent trainer/coach in educational settings and as a lecturer at the department of Developmental Psychology at the University of Groningen.

Henderien Steenbeek studied developmental psychology at the University of Groningen, where she defended her PhD thesis in 2006. The subject of her thesis was modeling dyadic child-peer interactions during play. She currently works as associate professor at the department of Developmental Psychology at the same university. In addition, she works as a professor at Teacher College at the University of Applied Studies in Groningen. In her work she focuses on child-adult and child-child interaction in play-educational settings. Her research themes are: children's learning processes in primary education (with a focus on children with special needs, and excellent performing children), and the dynamics of children's play. Most studies are focused on the observation of interaction behaviors in naturalistic circumstances. See http://www.rug.nl/staff/h.w.steenbeek/research.

Paul van Geert holds a doctoral degree from the University of Ghent (Belgium) and is a professor of developmental psychology at the University of Groningen in the Netherlands since 1985. He has had a pioneering role in the application of dynamic systems theory to a broad range of developmental areas, including early language development and second language acquisition; cognitive development in the context of learning-teaching processes; and social development including social interaction and identity. His main aim is to better understand the general nature of developmental dynamics, i.e. nature of the mechanism(s) that drive and shape a developmental process in an individual, as the individual, given his or her biological properties and potentialities interacts with his or her actively explored and transformed environment. He has been Fellow at the Center for Advanced Studies in the Behavioral Sciences and has held visiting professorships at the Unversities of Torino (Italy), Paris V and Reims (France), Trondheim (Norway) and Harvard University (Mind-Brain-Education programme). For his research and an overview of his artwork, see http://www.paulvangeert.nl.