Learning and Individual Differences 36 (2014) 101–109

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Learning through movement: A comparison of learning fraction skills on a digital playful learning environment with a sedentary computer-task

Baukje de Koning-Veenstra *, Henderien W. Steenbeek, Marijn W.G. van Dijk, Paul L.C. van Geert

Department of Developmental Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 9 September 2013
Received in revised form 18 August 2014
Accepted 24 October 2014

Keywords:
Physical activity
Arithmetic
Verbal interaction
Motivation
Case-study

ABSTRACT

Effects of physical exercise during educational tasks on learning are not obvious. This study examines the effects of movement on learning fraction skills at a physically active Playful Learning Environment (PLE). Employing a mixed-method approach, we investigated whether differences in motivational and verbal helping behaviour processes (underlying learning mechanisms) impacted learning gains. Results from 32 4th and 5th graders (n = 16 with the PLE, n = 16 in an equal sedentary computer task) approached significance, indicating that the PLE group showed higher learning gains compared to the SE group. Motivation was initially significantly higher for the PLE group. However, results indicated decreasing differences between the PLE and the sedentary group. There were no clear differences in the quality of verbal helping behaviour between the groups. Furthermore, subsequent utterances of verbal helping behaviour of two dyads selected in a case-study showed that their helping behaviour was sub-optimal for learning.

1. Introduction

The development of new powerful learning environments for engaging children in arithmetic activities in elementary education is a challenge for many educators. An example of a relatively new tool to teach children fraction skills is a digital Playful Learning Environment (PLE). This learning environment offers an active learning method, which aims to physically engage children in learning tasks, e.g. arithmetic, language or geography. With a PLE, children can create and play games on an outdoor playground while solving fraction problems. It aims to function as an interactive, playful learning environment through physical activity.

A PLE contains characteristics that aim to enhance powerful learning. Powerful learning can be seen as "learning that takes place effectively and efficiently, leading to permanent increases in relevant and usable knowledge and skills that stimulate and support further learning" (Veenstra, Van Geert, & Van der Meulen, 2011, p. 51). In powerful learning environments, learners are encouraged to construct their own knowledge and learn in realistic situations and together with others (De Jong & Pieters, 2006; Veenstra et al., 2011). Powerful learning environments stimulate active learning and constructive learning which are promoted by collaborative playful learning, and are aligned with individual differences, to reach high engagement (De Corte, Verschaffel, Entwistle, & Van Merriënboer, 2003; Hirsh-Pasek, Golinkoff, Berk, & Singer, 2009).

A PLE is based on the hypothesis that physical activity within the curriculum positively affects children's cognition, concentration and academic achievement (Donnelly & Lambourne, 2011; Donnelly et al., 2009). However, this type of physical activity can be defined in various ways. A distinction can be made between physical activity *during* and *after* the educational task (Trudeau & Shephard, 2008). This article is focused solely on physical activity *during* the educational task.

Research of Kangas (2010) and Kangas, Randolph, Ruokamo, and Hyvönen (2010) already indicates that the use of the PLE and various combinations of creative and playful learning methods is worthy of further research and implementation in educational practices. Their study shows, based on pre- and post test scores and qualitative results, that a PLE promotes academic achievement and creativity. However, it was not investigated whether the improvement was due to the specific PLE characteristics, since there was no control group involved in the study. Further, with only pre- and post-tests, valuable information about the learning processes diminishes (Flynn & Siegler, 2007; Steenbeek, Jansen, & Van Geert, 2012).

Effects of physical activity during the educational task are not obvious (Clinton, 2013). Some studies report that movement, whether it is during or outside the curriculum, has a positive effect on concentration, classroom behaviour, memory, self-esteem and reducing anxiety (Chomitz et al., 2009; Donnelly & Lambourne, 2011; Donnelly et al., 2009; Trudeau & Shephard, 2008). There are also studies with less convincing effects, which show that academic achievement neither increases, nor decreases through physical education (Ahamed et al., 2007; Carlson et al., 2008; Rasberry et al., 2010). Several studies report

* Corresponding author. Tel.: +31 623976303; fax: +31 503637602.
E-mail addresses: B.Veenstra@rug.nl (B. de Koning-Veenstra), H.W.Steenbeek@rug.nl (H.W. Steenbeek), M.W.G.van.Dijk@rug.nl (M.W.G. van Dijk), paul@vangeert.nl (P.L.C. van Geert).

http://dx.doi.org/10.1016/j.lindif.2014.10.002
1041-6080/

that physically active and fit children (whether this was due to physical activity within the curriculum or fitness in general) tend to have better academic achievement (Chomitz et al., 2009; Trost, 2007). However, a review of Taras (2005) reports that long-term improvement of academic achievement as a result of more vigorous physical activity (e.g. aerobic exercises or balance activities) is not well substantiated.

A possible explanation for the differences in outcomes of effects of physical activity might be that effects of physical activity can be either directly or indirectly linked to learning gains (Clinton, 2013). The direct physiological effect might be that this increases the flow of oxygen rich blood, water and glucose to the brain and the production of the mood-enhancing neurotransmitter dopamine, which in turn increases cognitive functioning (Galley, 2002; Hannaford, 1995, 2005). Brain scans show that children learn best when they are actually moving (e.g. jumping, swinging) and learning at the same time (Hannaford, 1995). The indirect effect might be that physical activity enhances on-task behaviour, compliance, attitudes towards learning, academic motivation, and attention, although these relationships seem to be small (Clinton, 2013). In addition, the quality of the task, teaching and school environment determine the amount of learning gains. However, it might be that children do not always experience a physical activity as engaging or intrinsically motivating. Children's expectancy-related beliefs (Eccles, Wigfield, & Schiefele, 1998) and subjective task values influence their motivation, including engagement, the amount of effort exerted, persistence and performance (Xiang, McBride, & Bruene, 2004). Therefore, a PLE should consist of characteristics that stimulate enjoyment, high expectancies of success, and an active role in children's learning. This will help children learn to value physically active tasks and foster authentic or intrinsic motivation. As a consequence, children will be more likely to engage in physically active tasks, which will have a positive effect on their learning (Ryan & Deci, 2000). High levels of engagement can affect the learner's attention, inquisitiveness and reflection (Price & Rogers, 2004). Optimal learning requires the process to be fun, which can increase children's engagement or authentic motivation (Hirsh-Pasek et al., 2009). As a consequence, deep learning can occur (Csikszentmihalyi, 1990; Ryan & Deci, 2000).

Effects on learning gains are dependent on multiple context factors: type of task (e.g. physically active or sedentary) and children and the quality of collaboration or interaction (e.g., Fuchs, Fuchs, Kazdan, & Allen, 1999; Gillies, 2004; Siegler & Alibali, 2005). One of the most consistent findings in the literature is the positive effect of high-quality verbal helping behaviour on learning gains (Gillies & Ashman, 1997; King, 2002; Topping, 2005; Webb & Farivar, 1994; Webb & Mastergeorge, 2003). Collaborative peer learning environments have potential for improving learning and increasing children's motivation, time on task and self-esteem (Webb & Mastergeorge, 2003), if collaboration consists of positive interactions that promote learning (Solomon, 1990). However, to ensure that collaborative learning is used effectively and to understand the effects of collaborative learning, interaction processes should be directly examined within collaborating dyads or groups (Solomon, 1990).

As known to us, no literature is available about the effect of physical activity on the quality of verbal interaction or collaboration. However, literature shows that the effect of the quality of verbal interaction on achievement might be mediated by motivational variables (Solomon, 1990). Since physically active tasks might have a positive effect on children's academic motivation and attitudes (Clinton, 2013), we expect a positive effect on the quality of verbal interaction during physical active tasks as compared to sedentary tasks. Therefore, in order to achieve optimal learning, it is important to develop a PLE that elicits high motivation and high-quality verbal helping behaviour between peers, which can potentially augment children's learning gains (e.g., Oortwijn, Boekaerts, Vedder, & Strijbos, 2008; Webb & Mastergeorge, 2003).

Research has already demonstrated that the computer can serve as a facilitator of social interaction (e.g., Mavrou, Lewis, & Douglas, 2010). Meta-analyses in secondary education have shown that students working with computers in small groups outperform students working with computers individually (Lou, 2004; Lou, Abrami, & d'Appolonia, 2001; Susman, 1998). However, learning benefits are promoted only when students are able to seek and give elaborated help effectively (Webb, Ing, Nemer, & Kersting, 2006).

In this study, we define high-quality verbal helping behaviour as utterances that are *beneficial* for learning, such as giving explanations, asking content related questions, and producing motivational utterances (Baker, D'Mello, Rodrigo, & Graesser, 2010; Oortwijn et al., 2008; Rojas-Drummond & Mercer, 2003). Low-quality verbal helping behaviour is defined as *unbeneficial* for learning, such as off-task utterances (Webb & Mastergeorge, 2003), and demotivating utterances (Baker et al., 2010). Furthermore, studies have shown that thinking aloud has a positive effect on student's performances, though not for the collaborating peer (Kotsopoulos, 2010; Siegler & Alibali, 2005).

To elicit high-quality verbal helping behaviour, it is important to develop tasks that scaffold children's knowledge and to select dyads working together with different levels of skills. Successful scaffolding can take place, which means that the high-ability peer (the scaffolder) scaffolds the lower ability peer (the scaffoldee) (Granott, 2005). Lower ability children can benefit from higher ability children, since the higher ability children can explain how to solve the fraction assignments. Higher ability children can also benefit from lower ability children, since the sharing knowledge with the lower ability peers have particular tutoring effects on the higher ability children (Bransford, Brown, & Cocking, 2000; O'Donnell, Hmelo-Silver, & Erkens, 2006; Vygotsky, 1978).

In this article, we will compare the effect of learning on a PLE to an equally sedentary computer task on children's fraction skills. The PLE condition contains all the elements of the sedentary computer condition plus the element of physical exercise. We expect that physical engagement creates an involvement and activeness in learning that a sedentary task does not.

In addition to focusing on the aspect of physical exercise, the following two important components of learning are taken into account in this study: 1) the frequency and quality of verbal helping behaviour between collaborating peers and 2) the various motivational aspects across time. We focus on the quality of verbal helping behaviour during the intervention, since utterances between a child and a peer mutually influence each other (Steenbeek & Van Geert, 2013). The mutual influences can give rise to high-quality verbal helping behaviour, which are associated with optimal functioning and learning, and over the long-term contribute to the forming of successful learning trajectories, or with ineffective learning, and over the long term contribute to the forming of unsuccessful learning trajectories (Nakamura & Csikszentmihalyi, 2002; Steenbeek & Van Geert, 2013). Therefore, studying patterns of subsequent verbal utterances between peers may generate useful insights into underlying mechanisms of the possible learning gains and gives information on successful or unsuccessful learning (e.g. Lavelli, Pantoja, Hsu, Messinger, & Fogel, 2004; Van Geert & Steenbeek, 2006).

To fully understand the underlying patterns of a learning process, insight is needed in the temporal unfolding of learning processes in *individual* children (Van Geert, 2009). Case study methodology is an appropriate way for studying individual trajectories (Flyvbjerg, 2006; Gerring, 2007; Yin, 2009). An individual case study can provide valuable information about possible underlying patterns of the learning process, even if they are not generalizable in the classical population-oriented sense of the word (Lee & Baskerville, 2003). A case study may contribute to our knowledge of the temporal interaction patterns in the form of which the learning takes place.

The first questions (1–3.1) in this study concern group analyses focusing on learning gains, and motivational and verbal helping behaviour in which we hope to see changes that might lead to the possible difference or correspondence in learning gains between the two groups. To answer the last question (3.2), we developed a case-study, which is a more detailed illustration of individual trajectories, thus, of what actually happened during playing the games. With this, the quality of help seeking and giving behaviour between collaborating lower and higher scoring children will be examined in more detail. We selected two dyads (a PLE and a sedentary dyad, both consisting of a higher and a lower scoring child) in which we expected to see relatively high-quality help seeking and giving behaviour, based on the theory of scaffolding.

1. Which tool (sedentary computer task (SE) versus playful learning environment (PLE)) is a more effective tool to teach fractions in children? We expect that
   - Learning gains in fraction skills of PLE > SE.
2. Is there a difference in motivation between PLE and SE? We expect that
   - Motivation of PLE > SE.
   - A gradual decline between PLE and SE across the weeks, since once the children will be familiar with the tasks, PLE motivation will slightly decrease, but will still be higher than SE motivation.
3. Is there a difference in frequency and quality of verbal helping behaviour between PLE and SE?
   3.1. On the group level, we expect that
- Verbal utterances of PLE > SE.
- PLE (number, quality) > SE (number, quality), with high-quality verbal helping behaviour (motivating utterances, questions, and explanations) and low-quality verbal helping behaviour (demotivating and off-task utterances).
   3.2. Analyses concern action–reaction chains (combinations of subsequent utterances) of two dyads. We expect that
- PLE dyad and SE dyad will show high-quality verbal helping behaviour.
- PLE dyad will show more and more diverse action–reaction chains than the SE dyad.
- PLE dyad will show more positive action–reaction chains than the SE dyad.
- SE dyad will show more negative action–reaction chains than the PLE dyad.

2. Methods

2.1. Participants

A total of 32 4th and 5th grade Dutch students of two elementary schools in the northern part of The Netherlands participated in this study: 16 children participated in the PLE group (PLE) and 16 children in the sedentary computer task group (SE). Children were assigned by the researcher to a condition by sample matching, so that pre-test scores on fraction skills, enthusiasm for arithmetic and physical activity were distributed about equally across the PLE and the SE (see Table 1). As a result, there were no significant differences between the PLE and SE for pre-test scores ($p = .25$), enthusiasm for arithmetic ($p = .46$) and enthusiasm for physical activity ($p = .46$). All children played the games in pairs of two children. Every week, the children played with a different child. Dyads within the conditions were randomly assigned, thus in the first PLE session it could be that a high scoring child had to play with a low scoring child, while in the same PLE session another high scoring child had to play with another high scoring child. All children played with both high and low scoring children (on the pre-test). At both schools, children played for 5 weeks on the PLE or SE task, however, at one school children played 10 sessions, at the other school children played 5 sessions. Since the number of children in the PLE group and SE group at both schools was equally divided (School 1: 8 SE, 8 PLE children; School 2: 8 SE, 8 PLE children), the differences in number of sessions should cause no limitations for the analyses. The number of children is relatively small, since we had to transcribe their verbal behaviours, which is very labour-intensive.

Table 1
Descriptive information of participants.

	PLE ($n = 16$)		SE ($n = 16$)	
Gender male/female	10/6		9/7	
	M	*SD*	*M*	*SD*
Age in months	127.2	9.19	125.6	9.6
Screening:				
Pre-test scores fraction skills	21.62	19.41	22.25	21.51
Enthusiasm arithmetic	2.75	.86	2.73	.88
Enthusiasm physical activity	3.38	.62	3.53	.52

2.2. Tools

2.2.1. Playful Learning Environment (PLE)

The study was conducted at playgrounds of two elementary schools. Before the study started, the children were already familiar with the PLE by using it in their school breaks and after school time. The PLE was not used yet for educational purposes. The PLEs at the schools consisted of an iStation, iGrid and iPosts and software that was provided by smartus.com. The iStation consists of a computer screen with audio. In this study, the iGrid was used, which is a pressure-sensitive jump mat and gives signals to the iStation. The fraction skill games that were created for this study and were implemented in the PLE were made by teachers and researchers on regular computers, based on a regular fraction skills learning method Ambrasoft (www.ambrasoft.nl) which is used in many elementary schools in The Netherlands, with the Smartus design software (see Fig. 1 for an example).

The child should jump on the 4/8 tile and then receives feedback with a sound whether it was the correct or incorrect answer. In case of giving an incorrect answer, the child has the opportunity to jump again on an alternative tile, until the correct answer is found. Each assignment consisted of nine answer possibilities (represented by nine tiles). The child had the opportunity to jump on eight different incorrect tiles until the correct answer appeared. Next, the following assignment appeared. In this study, there were 34 games available, with eight assignments per game. The difficulty of the games increased with game 1 as the easiest game and game 34 as the most difficult game. Collaborating dyads had to start in the next play session with the highest game number they had finished during the previous session.

2.2.2. Sedentary computer task

The SE group played the same fraction assignments on a regular computer in the classroom as the PLE group. Thus, the image, instruction, feedback and number of answer possibilities at the iStation screen were equal to the PLE group, but they did not have to jump on the correct answers, but had to click with the mouse on the correct answers.

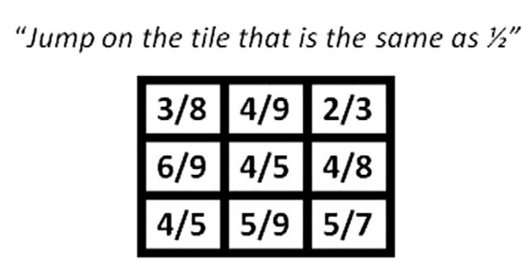

Fig. 1. Visual instruction on the iStation computer screen, an example of an assignment within the game that was developed.

2.3. Procedure

Five or ten sessions (dependent on the school) of approximately 20 min took place during morning hours at the elementary school during a five week period, with one or two sessions per week. The PLE children played the games at the outside playground and the SE children played the games at a computer in the classroom. A teacher and a researcher were available for questions about how to solve the fractions and for technical assistance.

2.4. Measures

2.4.1. Fraction skills

To measure learning gains in fraction skills, a pre- and a post-test of knowledge of fractions was administered. A paper-and-pencil assessment was composed with equal fraction assignments as in the regular teaching method of the children (Ambrasoft, www.ambrasoft.nl). The fractions in this test were a selection of the fractions, which were of increased difficulty, that were requested in the PLE and SE tasks. The children had 15 min to complete the test as far as possible. There were no children who finished the test. For each correct answer in the test, the children received one point, incorrect answers or empty answers were zero points. The total number of points indicates the level of knowledge of fraction skills.

2.4.2. Enthusiasm for arithmetic and physical activity

Enthusiasm for arithmetic and physical activity in general were assessed with a pre-test which consisted of a teacher questionnaire with which the teacher had to rate enthusiasm of each child. The raters had to categorize each individual child into the best fitting category (1–4, from no enthusiasm to high enthusiasm) for arithmetic as well as physical activity.

2.4.3. Motivation

To gain insight in task and situation-specific motivation while children engaged in the specific activities, motivation was measured with the Dutch version of the Online Motivation Questionnaire (OMQ) before the task started and after the task was finished in weeks 1, 3, and 5 (Boekaerts, 2002; Crombach, Boekaerts, & Voeten, 2003). This questionnaire assesses children's judgments of a learning situation in real classroom situations with a 19 item pre-test and a 10 item post-test, with mostly four response scales per item. It consists of multiple subscales: Mood, Task Anxiety, Self-efficacy, Success Expectancy, Task Attraction, Intended Effort, and Perceived Utility. A questionnaire, such as the OMQ, administered before and after completing an educational task, is a straightforward way of examining motivation (Crombach et al., 2003). The test–retest reliability ranged from .48 to .82, which can be considered as sufficient or good. The internal consistency was higher than .70 for all scales, except the result assessment, which was .61 (Boekaerts, 2002; Moos & Azevedo, 2008).

2.4.4. Verbal helping behaviour

In both groups, verbal utterances were video-taped with Inqscribe (version 2.2) (www.inqscribe.com), and coded for frequency and type (see Appendix A). Inter-observer reliability was determined by comparing two trained coders independently coding three play sessions (about 45 min in total) from the intervention group. The mean kappa coefficient was .814, which indicates highly reliable coders (Landis & Koch, 1977).

Explanations, questions and motivational utterances were considered as *beneficial* for learning (high-quality helping behaviour). Off-task and demotivating utterances were considered as *unbeneficial* for learning (low-quality helping behaviour). Thinking aloud was considered as *neutral* for learning. We also coded prompts, procedural remarks, and utterances about technology. Since we could not find literature about the effects of these utterances on learning gains, we consider these categories as *neutral* for learning.

2.5. Analyses

Differences in learning gains of fraction skills, motivation and verbal helping behaviour were analysed with statistical procedures consisting of descriptive analyses performed in Microsoft Excel and permutation tests using Monte Carlo analysis (Good, 1999; Todman & Dugard, 2001) performed in Poptools (Hood, 2008). The Monte Carlo analysis is particularly efficient to analyse small sample data that are not normally or regularly distributed, such as the data discussed in this study. This technique is used to simulate null hypotheses that relationships or properties, such as differences in learning gains or differences in motivation, are based on chance. Each significance test was based on 10,000 simulations. p-Values which were equal or smaller than .05 state whether hypotheses are supported or not. Effect sizes were calculated based on pooled standard deviations (Cohen, 1992).

In the case-study (research question 3.2), we selected two dyads (one PLE dyad and one SE dyad) with equal gender (a boy and a girl) which consisted of one high and one low scoring child with each about equal differences between pre-test scores on fraction skills, namely dyad 1 who showed 34 point difference and dyad 2 who showed 33 point difference between the higher and the lower scoring child within the dyad. To examine the quality of verbal helping behaviour, we analysed action–reaction chains, i.e. combinations of subsequent utterances. Based on columns of successive codes in Excel, we set up transition matrices, with which frequencies of action–reaction chains were calculated, based on percentages of the total number of combinations of subsequent utterances.

3. Results

3.1. Effect of PLE on knowledge of fractions

Table 2 shows the average scores, standard deviations and difference scores on pre- and post-test per group. The effect size of the learning gains was $d = .60$ ($p = .057$), which is a medium effect size (Cohen, 1992). This suggests approach significance in the expected direction in that the gains of the PLE group were almost significantly higher than the SE group. The difference scores of both groups were significantly larger than differences as expected on the basis of chance ($p < .001$). There are relatively big differences between the children in the knowledge of fraction skills within the groups, mainly in the pre-test, however, these differences between pre-test scores were not significant ($p = .467$).

3.2. Motivational processes

All subscale scores, for the PLE as well as for the SE are higher than 2.6 (scores can vary between 0 and 4) (Fig. 2). This means that children show on average relatively high motivation on all characteristics on all subscales. Further, the PLE group shows higher scores than the SE group in week 1 ($d = .43$; $p = .001$) and week 3, ($d = .39$; $p < .001$), which was a highly significant effect. Strikingly, no significant differences for week 5 were found ($d = .19$; $p = .115$), since the SE group

Table 2
Descriptive information of pre- and post-test scores of fraction skills of the PLE and sedentary groups.

	PLE		Sedentary	
	M	*SD*	*M*	*SD*
Pre-test ($n = 16$)	21.62	19.41	22.25	21.51
Post-test ($n = 14$)	56.14	24.68	42.50	27.97
Difference score (post- minus pre-test)	34.52		20.25	

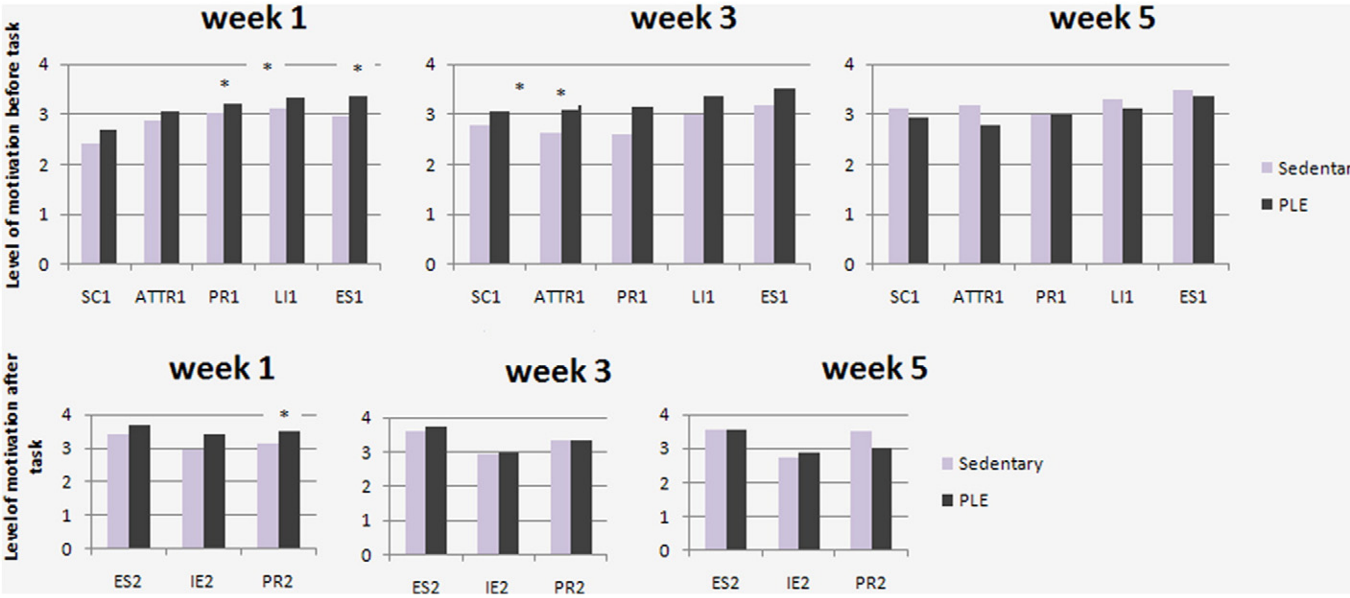

Fig. 2. Average scores on subscales on the Online-Motivation Questionnaire per week before (1) and after (2) the task; abbreviations of subscales: SC1 = subjective competency; ATTR1 = task attractiveness; PR1 = personal relevance; LI1 = learning intention; ES1 = emotional state before; ES2 = emotional state afterwards; IE2 = intended effort; PR2 = perceived results (* = $p < .05$).

shows somewhat higher motivation scores than the PLE group on almost all subscales, which contradicts expected differences.

3.3. Verbal helping behaviour between dyads

3.3.1. An overall overview

Contrary to our expectations, SE children showed a higher number of utterances ($d = .30$; $p = .068$), which approached significance. The average amount of verbal utterances per minute, of the PLE was $M = 8.12$; $SD = 4.06$ and of the SE was $M = 9.40$; $SD = 4.59$.

Based on our expectation that PLE children show more high-quality helping behaviour than SE children, we analysed the frequencies of high versus low-quality helping behaviour in comparison to the total amount of utterances in both groups (Fig. 3).

Mostly, children in both groups gave prompts (P) (about 20% and 27% of the total number of verbal utterances), and showed off-task verbal utterances (O) (about 12% and 13% of the total number of verbal utterances). The biggest part of utterances was coded as 'procedural category' (PC) (about 30% and 27% of the total number of verbal utterances), which was task-related. A relatively high number of utterances was incomprehensible (I), due to background noise (14%). There is no reason why background noise should have suppressed a particular type of utterance.

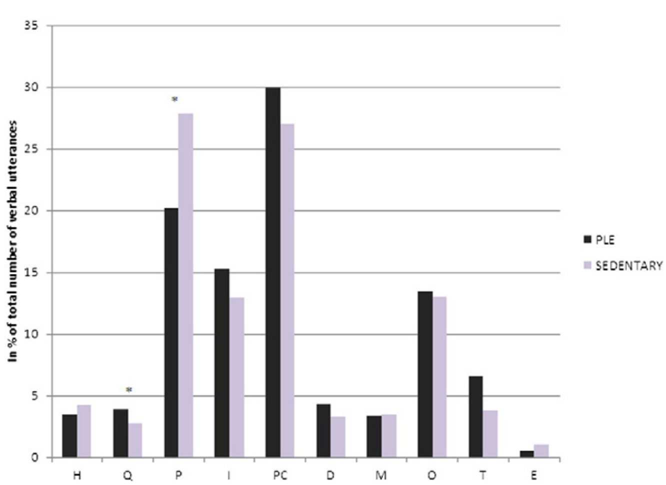

Fig. 3. Percentage of each category of interaction of the total amount of interaction; abbreviations of subscales: H = thinking aloud; Q = questioning; P = prompting; I = incomprehensible; PC = procedural category; D = demotivating; M = motivating; O = off-task behaviour; T = technique; E = explaining (* = $p < .05$).

The differences between verbal utterances between the two groups are small. Differences were only significant for prompts and questions, that is, the SE group showed more prompts ($p = .005$) and the PLE group showed more questions ($p = .029$). Prompts were mainly in the form of giving (in)correct answers, mainly by mentioning numbers. Children were mainly questioning in the form of short sentences, such as "Why?" or "Where?". No significant differences were found in the number of motivating comments, off-task comments, and explaining comments between the PLE and SE groups.

3.3.2. Case study: a comparison of the quality of helping behaviour in dyads

The PLE dyad showed more verbal reactions ($n = 5.31$ utterances per minute) than the SE dyad ($n = 3.37$ utterances per minute), while the average scores showed the opposite (see Section 3.3.1). If not counting the incomprehensible utterances (I) of the PLE dyad, there are about an equal number of combinations of subsequent utterances (action-reaction chains) in the SE ($n = 29$) and the PLE dyad ($n = 26$).

The action–reaction chains (the "arrows" in Figs. 4 and 5) consist of utterances of the first child and reactions of the other child, or utterances of the first child after which the same child responses to him/herself (self-iteration). With these action–reaction chains, insight can be obtained in common patterns within verbal helping behaviour, e.g., is a question regularly followed by an explanation (see Figs. 4 and 5).

Results show that the PLE dyad shows more positive utterances, such as questions (Q), while the SE dyad shows more explanations (E) (Figs. 4 and 5). However, the SE dyad also shows more negative action–reaction chains, with more demotivating (D) and off-task remarks (O). No clear differences were found in motivating remarks (M). As compared to the group analysis (Fig. 3), the structures of interactions of the two dyads are not so different, with also relatively many prompts (P) and procedural category (PC) remarks and few questions (Q), explanations (E) and motivating remarks (M).

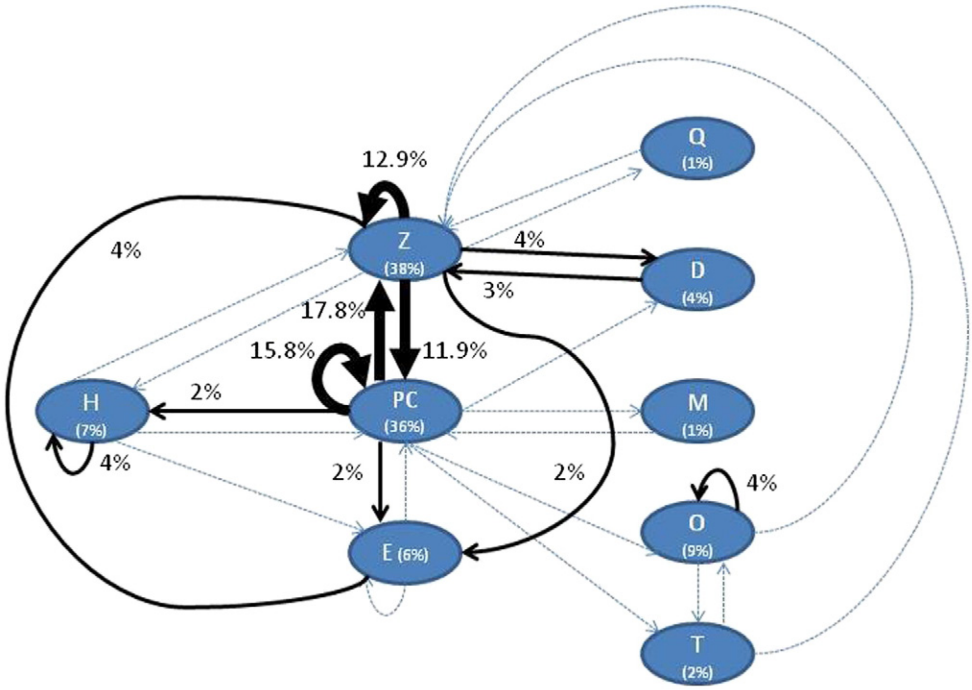

Fig. 4. Transition diagram of the SE dyad. The dashed arrows represent chains of 1%. The dashed arrows represent chains corresponding with a transition probability of 1%. The thick arrows represent the main chains, corresponding with a transition probability above 10%. The normal arrows represent chains corresponding with a transition probability between 2% and 10%.

4. Discussion

In general, the analysis approached significance ($p = .057$), indicating that a PLE seems to be a more effective learning tool as compared to the SE task ($d = .60$). Not only the learning gains of both groups indicate that the PLE is more effective than the SE, but also motivation of both groups across time indicates that the PLE is a more motivating tool than the regular computer. However, there was a clear indication of decreasing differences between motivation scores between the PLE and the SE group across time. Furthermore, the SE group seems to elicit more verbal utterances than the PLE group ($d = .3$; $p = .068$). When we take a look at the categories of verbal utterances, there were almost no differences between both groups. It seems that a physical activity as presented in this study did not have as much influence on the quality of verbal helping behaviour as was expected.

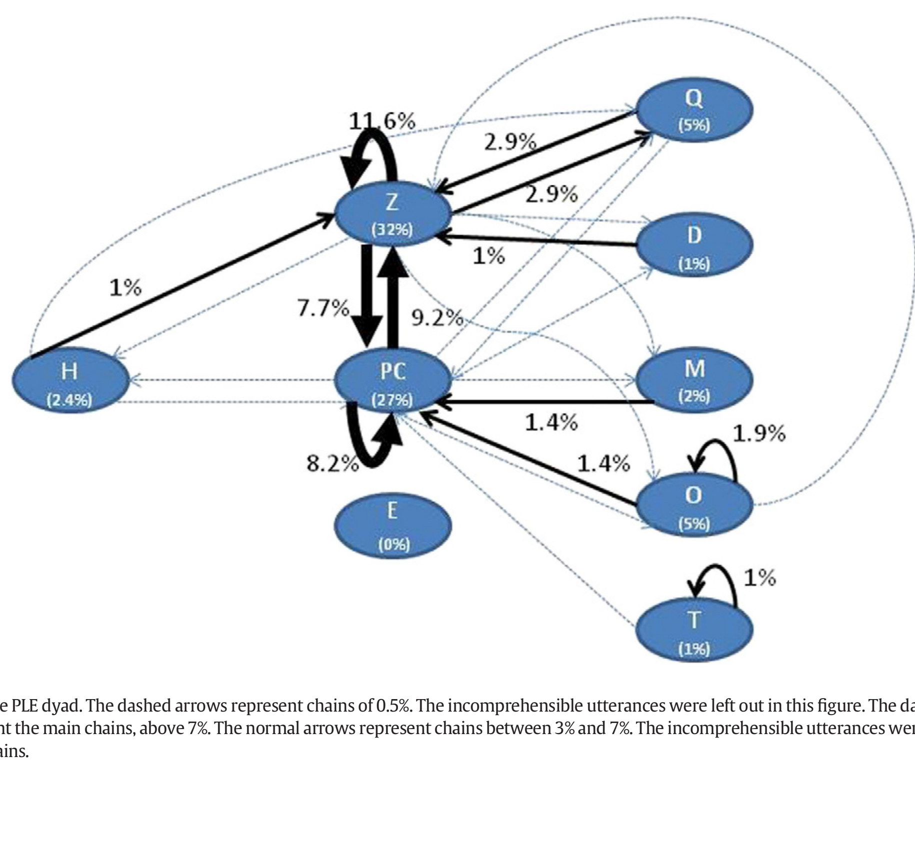

Fig. 5. Transition diagram of the PLE dyad. The dashed arrows represent chains of 0.5%. The incomprehensible utterances were left out in this figure. The dashed arrows represent chains of 0.5%. The thick arrows represent the main chains, above 7%. The normal arrows represent chains between 3% and 7%. The incomprehensible utterances were left out in this figure. The thick arrows represent the main chains.

In this study the sample is relatively small, given the labour-intensive nature of the analyses. The learning-effect is nevertheless quite considerable ($d = .6$), and the p-values are slightly bigger than .05, which is quite acceptable (Gliner, Leech, & Morgan, 2002; Tyler, in Daniel, 1998, p. 24). Therefore, we pose that the PLE can be a more powerful learning tool for children as compared to the regular computer, at least in certain tasks such as fraction skills.

The analyses of the case-study have shown that both dyads showed short utterances and that the PLE dyad showed more, but not more diverse, verbal utterances than the SE dyad, which was contrary to the group analyses. This indicates that averages should not be identified with properties that apply to every child of which the average has been computed. Further, the results show that the PLE dyad showed more positive utterances and the SE dyad showed more negative action–reaction chains. However, the collaboration between dyads was relatively superficial, while we expected to see high-quality verbal help seeking behaviour. The frequency of verbal utterances was higher in the PLE dyad. However, action–reaction chains suggest that the type of collaboration was about equal between both dyads. The differences between both dyads were relatively small and not entirely in line with the hypotheses. It might also mean that differences between children (e.g. personality) or natural fluctuations from session to session could be more important than the difference in task affordances.

In this study, more classical analyses were combined with process analyses, to gain insight into the process of learning, before, after and during fulfilling the task itself. All results indicate that this game, both at the PLE and SE, neither does stimulate children to give (extensive) explanations, nor intensive collaboration between dyads. This study confirms that collaboration between dyads is easily established, but it is not guaranteed that optimal learning is taking place (Webb & Cox, 2004). It might be cautiously concluded that higher learning gains might be achieved by teaching the children how to improve their mutual helping behaviours, than by introducing a physical element in their learning.

However, this study has some clear limitations. The small difference between the PLE and the SE group in motivation and quality of verbal helping behaviour might have been caused by the type of game presented in the study and technical problems with the grids on the jump mat. If the PLE game would have been more physically demanding and motivating (e.g. by introducing a stronger competition element) than the current PLE game, the learning gains might have been more considerable. However, this is mere speculation, since increase in physical demands and the focus on the competition element could equally well lead to lesser instead of higher learning games.

Relatively many utterances of the PLE dyad were incomprehensible (about 14%), which might reduce the statistical power of the comparison, although we do not expect that important verbal utterances were more suppressed than unimportant ones. Another possible limitation of this study might be that the collaborating dyads changed every week, to correct for possible learning effects of working with the same dyad. This might have been problematic for some children, since they had to start with a 'new' collaborating peer every week. Further, collaborating dyads had to start in the next play session with the highest game number they had finished during the previous session, which did not always correspond to the skill level of one child within the dyad. However, if the dyads did not change, it possibly greatly hampered those dyads that did not really like each other.

Based on the results of this study, we do not expect that physical activity on the PLE solely enhances learning gains. We expect that with more adaptive task and teacher scaffolding, which is also focused on the interactive process of helping behaviour, learning gains will increase (Kollar, Fischer, & Hesse, 2006). With challenge, variation and competition within the games, thus with more playful learning (Hirsh-Pasek et al., 2009), we expect that children will be more engaged to play educational games (Veenstra et al., 2011). In addition, we expect that the probability will be higher that deep learning will take place in more children, mainly in the PLE condition, which might result in more successful learning trajectories. However, the risk of a competitive game is that children will explain almost nothing to each other in order to win the games. Therefore, a balance should be found between physical challenge, competition, helping behaviour and teacher's scaffolding during playing the games. As a consequence, learning gains in children playing with the physically active PLE will be optimally enhanced.

Acknowledgements

Preparation of this manuscript was made possible by an award from Platform Bèta Techniek, The Hague and elementary school 't Vrieske Honk, Vriescheloo and OBS Angelslo, Emmen. Further, we would like to thank Master's students Hylke Faber and Alev Cumert for transcribing and coding videotapes.

Appendix A. Coding scheme "Arithmetic on the PLE and the sedentary computer" (Veenstra & Faber, 2011)

General points before

1. First watch the whole video once (without stopping or pausing). The video can be watched in InqScribe.
2. Then you start with transcribing. If a child says something, you click on 'stop'; you have to write down (transcribe) what the child says (with the name of the child in front) in InqScribe. With this, it will be clear who says what and when. If a child says something, but is incomprehensible, you have to transcribe "…".
3. After transcribing, you have to give every sentence a code in Excel. The codes are described at the next page. If the video is already transcribed, then you have to code the sentence with the help of the videos; this is because the context and intonation will be clearer. You cannot code without the videos.
4. If you are not sure about a code, watch that part of the video again.

Operationalisations of the codes

Prompting or reading what is on the screen (P)

Reading or giving prompts as displayed on the screen of the PLE or sedentary computer. If the child only mentions a number, this will be coded as (P). This can be an utterance of both children within the dyad. Giving prompts exists in giving an answer on the question without giving explanations, e.g. only a number. It can also exist in telling which answer is correct or which tile or key should be pressed. Also in case of giving a wrong answer, this should be coded as a 'P'. The most important element is that there is no given explanation. Remarks such as "that one" or "there" or "these" also belong to giving prompts.

Thinking aloud (H)

With thinking aloud, thinking steps of the solving process will be spoken aloud. Repeating the questions also falls into this category, as well as if a child says what he or she intended to do, such as "I will jump on this tile". If a child mentions just one number, it should not be coded as an 'H', since it does not reflect a process or way of acting/thinking. If a child counts for example parts of a pie within the game, this should be coded as an 'H'. More examples of 'H' are:

> "You have 12 parts, and then you share that with 6 children …""The whole pie was 10 euros, then one part is …""One, two, three, four …"

Giving explanations (E)

With giving explanations, explanations about the content of the assignment will be given. This should regularly be followed by a question. Explanations can be expressed by splitting the question in small pieces or by reformulating a question, so that the questioner can solve the task. Giving explanations about the functions of the tool (PLE or computer) does not belong to this category, but has to be coded with a 'T' (Technology). Examples of explanations are:

> "Then you have to multiply three"."That is two times five, thus ten"."There is one coloured, so nine are left".

Questioning (Q)

All questions that are focused on solving the assignment on the screen should be coded as a 'Q'. Questions about technological aspects or off-task questions do not belong to this category. Questions about reading the assignment should be coded as a 'Q'. Further examples of a 'Q' are:

"What was the assignment/question?""Do you want to read the question?""Shall I do this?""I'm doing great isn't it?""Shall I jump now?""Is that the same as 1/3rd?""3/4th?""Which assignment are you solving now?"

Motivating (M)

In this category utterances which one or both children stimulate to proceed with solving the assignments belong. Examples are:

"Great!""It's a pity"."Go again"/"Do again"/"Go ahead"."Yes!"

Demotivating (D)

All utterances of children/both children that don't stimulate to proceed with solving the assignments, utterances that have a demotivating intonation. Examples are:

"Don't you know that?!""You are doing badly, I was much further".

Remarks that reflect irritations about collaborating with another child, e.g.:

"Stop that!""Don't say anything …""Go away"."I don't like this assignment"."The assignment is too long".

Technology (T)

Utterances concerning technological functioning of the PLE or the computer belong to this category. This can be questions as answers about the technological side of the PLE.

Examples are questions to assist with pressing a tile, remarks about resetting the screen or shortcomings of a tile. Remarks about shortcomings of the PLE or computer also belong to this task, such as "I can't see it, go away". If remarks will be made which are not specifically about the functioning of the tool, these should be coded with a 'D', 'M' or a 'PC', such as: "I find this stupid": which has to be coded as a 'D'.

Procedural category (PC)

In this category utterances which cannot be placed into another category, are more procedural utterances and cannot be coded as 'off-task' belong. In this category utterances which are related to the task, but do not definitely contribute to the task belong, e.g.:

"Oh" and "Oops"."Yes"/"No"/"Not"."Name of the child"."We only have to play for just one minute".

Incomprehensible (I)

If an utterance is incomprehensible, this has to be coded as an 'I'. This is only the case if in the transcript "…" is transcribed.

Off-task behaviour (O)

Every question or remark which does not have anything to do with the task, such as other activities. If a child says that he/she does not want to play anymore, this has to be coded as 'D'.

References

Ahamed, Y., Macdonald, H., Reed, K., Naylor, P. -J., Liu-Ambrose, T., & McKay, H. (2007). School-based physical activity does not compromise children's academic performance. *Medicine & Science in Sports & Exercise*, *39*, 371–376.

Baker, R. S. J. d, D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241.

Boekaerts, M. (2002). The online-motivation questionnaire: A self-report instrument to assess students' context sensitivity. In P. Pintrich, & M. L. Maehr (Eds.), *New directions in measures and methods* (pp. 77–120). Oxford, UK: Elsevier Science, Ltd.

Bransford, J. D., Brown, A. L., & Cocking, R. (2000). *How people learn*. Washington DC: National Academy Press.

Carlson, S. A., Fulton, J. E., Lee, S. M., Maynard, L. M., Brown, D. R., Kohl, H. W., et al. (2008). Physical education and academic achievement in elementary school: Data from the early childhood longitudinal study. *American Journal of Public Health*, *98*(4), 721–727.

Chomitz, V. R., Slining, M. M., McGowan, R. J., Mitchell, S. E., Dawson, G. F., & Hacker, K. A. (2009). Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the northeastern United States. *The Journal of School Health*, *79*(1), 30–37.

Clinton, J. (2013). Physical activity. In D. Hattie, & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 33–35). New York: Routledge.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Crombach, M. J., Boekaerts, M., & Voeten, M. J. M. (2003). Online measurement of appraisals of students faced with curricular tasks. *Educational and Psychological Measurement*, *63*(1), 96–111.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: HarperCollins.

Daniel, L. G. (1998). Statistical Significance Testing: A historical overview of misuse and misinterpretation with implications for the editorial policies of educational journals. *Research in the Schools*, *5*(2), 23–32.

De Corte, E., Verschaffel, L., Entwistle, N., & Van Merriënboer, J. (2003). *Powerful learning environments: Unravelling basic components and dimensions*. Amsterdam: Pergamon.

De Jong, J., & Pieters, J. (2006). The design of powerful learning environments. In P. A. Alexander, & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 739–754) (2nd ed). Mahwah: NJ: Lawrence Erlbaum Associates.

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., et al. (2009). Physical Activity Across the Curriculum (PAAC): a randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine*, *49*(4), 336–341.

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine*, *52*(1), S36–S42.

Eccles, J.S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.). Handbook of child psychology (5th ed., pp. 1017–1095). New York: Wiley.

Flynn, E., & Siegler, R. (2007). Measuring change: Current trends and future directions in microgenetic research. *Infant and Child Development*, *16*(1), 135–149.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*, 219–245.

Fuchs, L. S., Fuchs, D., Kazdan, S., & Allen, S. (1999). Effects of peer-assisted learning strategies in reading with and without training in elaborated help giving. *The Elementary School Journal*, *99*(3), 201–219.

Galley, M. (2002). Texas requires elementary schools to offer 2-plus hours of physical education. *Education Week*, *21*(29), 10–13.

Gerring, J. (2007). *Case-study research: Principles and practices*. Cambridge: Cambridge University Press.

Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction*, *14*(2), 197–213.

Gillies, R. M., & Ashman, A. F. (1997). Children's cooperative behavior and interactions in trained and untrained work groups in regular classrooms. *Journal of School Psychology*, *35*, 261–279.

Gliner, J. A., Leech, N. L., & Morgan, G. A. (2002). Problems With Null Hypothesis Significance Testing (NHST): What do the textbooks say? *The Journal of Experimental Education*, *71*(1), 83–92.

Good, P. I. (1999). *Resampling methods: A practical guide to data analysis*. Boston: Birkhauser.

Granott, N. (2005). Scaffolding dynamically toward change: Previous and new perspectives. *New Ideas in Psychology*, *23*, 140–151.

Hannaford, C. (1995). *Smart moves: Why learning is not all in your head*. Arlington, VA: Great Ocean.

Hannaford, C. (2005). *Smart moves*. Arlington, VA: Great Ocean Publishers.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). *A mandate for playful learning in preschool*. New York: Oxford University Press.

Hood, G. (2008). *Poptools [computer software]*. Canberra, Australia: Pest Animal Control Co-operative Research Center (CSIRO).

Kangas, M. (2010). Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking Skills and Creativity*, *5*(1), 1–15.

Kangas, M., Randolph, J., Ruokamo, H., & Hyvönen, P. (2010). An international investigation into playful learning environments and academic achievement. *Paper presented at Annual Meeting of the American Educational Research Association (AERA), April 30–May 4, 2010 Conference. USA, Denver*.

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory Into Practice*, *41*, 33–39.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts — A conceptual analysis. *Educational Psychology Review*, *18*(2), 159–185.

Kotsopoulos, D. (2010). An analysis of talking aloud during peer collaborations in mathematics. *International Journal of Science and Mathematics Education*, *8*(6), 1049–1070.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lavelli, M., Pantoja, A. P. F., Hsu, H. C., Messinger, D., & Fogel, A. (2004). Using microgenetic designs to study change processes. In D. G. Teti (Ed.), *Handbook of research methods in developmental science*. Cambridge: Blackwell Publishers.

Lee, A., & Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research*, *14*(3), 221–243.

Lou, Y. (2004). Understanding process and affective factors in small group versus individual learning with technology. *Journal of Educational Computing Research*, *31*(4), 337–369.

Lou, Y., Abrami, P. C., & d'Appolonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449–521.

Mavrou, K., Lewis, A., & Douglas, G. (2010). Researching computer-based collaborative learning in inclusive classrooms in Cyprus: The role of the computer in pupils' interaction. *British Journal of Educational Technology*, *41*(3), 486–501.

Moos, D., & Azevedo, R. (2008). Exploring the fluctuation of motivation and use of self-regulatory processes during learning with hypermedia. *Instructional Science*, *37*, 203–231.

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 89–105). New York, US: Oxford University Press.

O'Donnell, A. M., Hmelo-Silver, C., & Erkens, G. (Eds.). (2006). *Collaborative learning, reasoning, and technology*. Mahwah, NJ: Lawrence Erlbaum.

Oortwijn, M. B., Boekaerts, M., Vedder, P., & Strijbos, J. -W. (2008). Helping behaviour during cooperative learning and learning gains: The role of the teacher and of pupils' prior knowledge and ethnic background. *Learning and Instruction*, *18*(2), 146–159.

Price, S., & Rogers, Y. (2004). Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education*, *43*(1–2), 137–151.

Rasberry, C. N., Lee, S. M., Robin, L., Laris, B. A., Russell, L. A., Coyle, K. K., et al. (2010). The association between school-based physical activity, including physical education, and academic performance: A systematic review of the literature. *Preventive Medicine*, *52*(1), 1–129.

Rojas-Drummond, S., & Mercer, N. (2003). Scaffolding the development of effective collaboration and learning. *International Journal of Educational Research*, *39*, 99–111.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Siegler, R. S., & Alibali, M. W. (2005). *Children's thinking.* Upper Saddle River, NJ: Prentice Hall.

Solomon, D. (1990, August 10–14). Cooperative learning, intragroup dynamics, and student outcomes. *Paper presented at the Annual Meeting of the American Psychological Association (98th, Boston, MA)*.

Steenbeek, H. W., Jansen, L., & Van Geert, P. L. C. (2012). Scaffolding dynamics and the emergence of problematic learning trajectories. *Learning and Individual Differences*, *22*(1), 64–75.

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*(2), 233–267.

Susman, E. B. (1998). Cooperative learning: A review of factors that increase the effectiveness of cooperative computer-based instruction. *Journal of Educational Computing Research*, *18*(4), 303–322.

Taras, H. (2005). Physical activity and student performance at school. *Journal of School Health*, *75*(6), 214–218.

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

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, *25*, 631–645.

Trost, S. G. (2007). *Active education: Physical education, physical activity and academic performance (research brief).* San Diego, CA: Robert Wood Johnson Foundation Active Living Research.

Trudeau, F., & Shephard, R. J. (2008). Physical education, school physical activity, school sports and academic performance. *International Journal of Behavioral Nutrition and Physical Activity*, *5*, 10.

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*. New York: Cambridge University Press.

Van Geert, P. L. C., & Steenbeek, H. W. (2006). The dynamics of scaffolding. *New Ideas in Psychology*, *23*, 115–128.

Veenstra, B., & Faber, H. (2011). *Coding scheme "Arithmetic on the PLE and sedentary computer".* University of Groningen (Unpublished internal document).

Veenstra, B., Van Geert, P. L. C., & Van der Meulen, B. F. (2011). Is edutainment software really educational? A feature analysis of Dutch edutainment software for young children. *Netherlands Journal of Psychology*, *66*(2), 50–67.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Webb, M., & Cox, M. (2004). A review of pedagogy related to information and communication technology. *Technology, Pedagogy and Education*, *13*(3), 235–286.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal*, *31*, 369–395.

Webb, N. M., Ing, M., Nemer, K. M., & Kersting, N. (2006). Help seeking in cooperative learning groups. In R. S. Newman, & S. A. Karabenick (Eds.), *Help seeking in academic settings: Goals, groups and contexts* (pp. 45–88). Erlbaum.

Webb, N. M., & Mastergeorge, A. (2003). The development of students' learning in peer-directed small groups. *Cognition and Instruction*, *21*, 361–428.

Xiang, P., McBride, R. E., & Bruene, A. (2004). Fourth graders' motivation in an elementary physical education running program. *The Elementary School Journal*, *104*(3), 253–266.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, California: Sage Publications.