Utrecht University

A multilevel study into the effect of tablet use in Dutch primary schools on reading achievements

Bachelor thesis

Author: Feline Wafelaar [6173233]

Date: 15 – 06 – 2020

Supervisor: Kim Stienstra

Field of study: Sociology, Utrecht University

Abstract

This study utilized data from the 2016 Progress in International Reading Literacy Study to examine the effect of tablet usage in schools on reading achievements of children. It further addressed the gender gap in achievements by exploring whether gender moderates the relationship between tablet use and achievements. The sample consisted of 2989 fourth grade students from Dutch primary schools. A multilevel model was conducted in which both student- and school-level predictors were included. Contrary to the positive effect of tablet use on reading achievements found by most previous studies, the findings suggested no relationship between tablet use and achievements. There was also no evidence found that gender moderates this relationship. Results are critically discussed from a methodological and theoretical perspective, and recommendations for future research are being made.

Keywords: tablets, reading achievements, primary school, gender, moderator, PIRLS, multilevel model

Introduction

Ever since the beginning of the 1980s educational institutions started to implement technological devices into their school system (Sharples, Taylor & Vavoula, 2007). Worldwide, an increasing number of digital devices become available every day, which can be used for educational purposes. Many schools have now replaced their traditional textbooks for tablets due to their affordable prices and strong battery power (Timmermann, 2010; Haßler, Major & Hennessy, 2016; Fokides & Atsikpasi, 2017). This ranges from adopting the 'Bring Your Own Device' programme, in which parents are encouraged to buy tablets for their offspring (Irwin & Jones, 2014), to schools equipping each student with a tablet (Jones, 2014). The take up of tablets in primary schools is also gaining popularity in the Netherlands, where schools use them in addition to their traditional teaching methods, or as replacement of it (Faber & Visscher, 2016). In fact, the Netherlands is at the leading edge of the digitization of education in Europe (Beblavý et al., 2019). In contrast to other countries, it has highly developed institutions and policies on digital learning, and great availability and usage of digital learning materials (Beblavý et al., 2019).

As the use of technological devices in educational settings grows, research tries to unravel how these technologies affect achievements (McFarlane, Triggs & Wan, 2008). Most studies into the effect of tablets on school achievements were about mathematical abilities (e.g. Outhwaite, Gulliford & Pitchford, 2017; Papadakis, Kalogiannakis & Zaranis, 2016), whereas research on reading skills is scarce. Studies that did investigate reading achievements only looked at the use of tablets in general (i.e. in- and outside of schools), only took into account low-achieving groups, or focused on experiences instead of achievements. Examining tablets in combination with reading performance is particularly interesting, because tablets may take away the ‘boring image’ that exists among primary school pupils about reading.

The use of tablets in schools can have both positive and negative effects on the reading performance of children. One advantage of tablets is that they allow for personalized learning (Shuler, Winters & West, 2012). That is, the difficulty of the assignment is based on students’ comprehension of the material. With traditional teaching methods all students have to do the same assignments, regardless of their level of comprehension. What is more, the interactive and playful way of learning increases motivation and dedication among children (Furió et al., 2015). Additionally, research suggests that boys will profit more from the use of tablets than girls, because it better fits their learning style and they are more confident with tablets (Carrier, 2009; Vekiri & Chronaki, 2008). Since girls in general outperform boys in reading skills (Lynn & Mikk, 2009), tablets might be useful tool to reduce this gap. That is, if boys gain more from the use of tablets than girls, this consequently compensates for the gender inequality in reading skills. Tablets can also negatively influence performance because children are more easily distracted when they use tablets (Sheppard, 2011). Further, although tablets engage children because they learn in a way that comes naturally (i.e. by trial and error), they can also get frustrated and bored when they do not succeed (Cohen, Hadley & Frank, 2012). Moreover, teachers often lack the skills that are needed to make the use of tablets in educational settings successful (Van de Bogart, 2012).

The discussion above illustrates the different theoretical ideas that exist about the positive and negative effects of tablets. Scientific research does not provide a definite answer either because results are fragmented (Krumsvik, Berrum & Jones, 2018; Churchill, 2012; Hutchison, 2012). Although most studies detect a positive effect of the use of tablets on reading achievements (e.g. Korat, 2010; Harmon, 2011; Flewitt, Messer & Kucirkova, 2015) there are also some studies that find no difference between using tablets or printed texts (e.g. Huang et al., 2012; Dundar & Akcayir, 2012) or even a negative effect (e.g. Gasparini & Culén, 2012; Sheppard, 2011). In addition to these studies, an overarching literature review reports positive effects of tablets on school achievements in general but note that it is often unclear if these positive effects can be attributed to the tablets (Haßler, Major & Hennessy, 2015). The authors state that most evidence lacks generalizability, and explanations for why and how tablet usage improves learning outcomes remains puzzling. In sum, these findings suggest that tablets positively influence reading achievements, but more research is needed in order to draw firm conclusions about the efficacy of tablets.

This research aims to investigate how the use of tablets in Dutch primary schools affects the reading achievements of children. In addition to contributing to the gap in scientific knowledge, it also ties in with social developments in the Netherlands. In recent years, the use of tablets in Dutch primary schools have caused controversy. The opening of a number of iPad schools by Maurice de Hond in 2013, in which the iPad was used as primary teaching tool, mainly sparked this discussion. Dutch newspapers published articles about children getting distracted easily, the difficult role of teachers, but also how tablets might enhance learning (e.g. NRC, 2019). Since the iPad schools went bankrupt in 2018, both teachers and parents remain sceptic towards the use of tablets in schools.

On top of studying the effect of tablet use on reading achievements, I will examine to what extent the effect of tablets on achievements differs between boys and girls. Insights into the possible moderating role of gender will expose whether tablets compensate for gender inequality in reading achievements. This leads to valuable information about gender inequality in reading achievements, which in turn is of use for the design of policy interventions. It is important to find ways to reduce the gender gap so that highest economic and social return on education can be harvested (OECD, 2012). In sum, the research question I intend to answer is two-folded:

1. *What is the effect of the use of tablets for schoolwork in Dutch primary schools on the reading achievements of children?*
2. *To what extent does the effect of the use of tablets for schoolwork in Dutch primary schools on reading achievements differ between boys and girls?*

The data that will be used is from the Progress In International Reading Literacy Study (PIRLS), a study that assesses the reading achievements of young children in their fourth year of schooling (group six in Dutch) in 61 countries. This is a unique and highly relevant dataset because it contains extensive data on reading achievements, and extensive data on school characteristics including tablet use. Also, only little research into tablets has been carried out in primary school settings (McPhee, Marks & Marks, 2013). For the aims of this study I utilize the Dutch subsample; a representative sample that followed 4206 students from 226 schools in the Netherlands in 2016 (PIRLS, 2016). Other countries are deliberately left out because in contrast to the Netherlands, most countries never - or only sporadically - use tablets in schools.

Including these countries would then lead to biased results. The Netherlands in itself is an interesting case to look at since it is on the leading edge of tablet use, and no relevant research into the effects of tablet on achievements has been conducted here yet. Only Faber & Visscher (2016) and Molenaar, Campen & Gorp (2016) have used the Dutch context so far. They investigated the impact of the app 'Snappet', an app with which children learn to spell and count, on the achievements of children. However, their outcomes remain debatable since it lacks generalizability, it was financed by the developer of the application, and it only looks at this specific application.

This study further sheds new light on the effect of tablets on reading abilities since it, in contrast to most of the above studies, employs a quantitative research design with a large representative study sample. Other differences exist in terms of tablet use and reading achievements. I will analyse the use of tablets for school purposes in general, instead of merely the use of a particular type of tablet or software program on it as done by previous studies. Also, I will look at tablets that are used for schoolwork in general and not those exclusively used for reading. At last, reading achievements are broadly measured by looking at their comprehension of words and texts, vocabulary and reading ability, whereas other scholars studied each aspect independently.

Theory

Constructivism

In order to estimate how the usage of tablets affects the reading achievements of children, one needs to understand the learning processes of children. In the past, school practices were mainly based on a behaviouristic perspective on learning (Druin & Solomon, 1996). This perspective encompasses the idea that teachers provide children with information, who are just simply copying the knowledge given by their teachers. One of the teaching styles that supports the notion of behaviourism is the drill-and-practice method (Suppes, 1980 as cited in Druin & Solomon, 1996). This instructional method focuses on the systematic repetition of specific skills (e.g. subtraction or spelling) (Lim, Tang & Kor, 2012). Downfalls of this teaching method are that children become bored easily (Druin & Solomon, 1996) and only memorize the material instead of truly understanding it (Vosniadou, 2001).

Due to these disadvantages, the behaviouristic perspective gave the floor to cognitivism in the 1970s (Tangdhanakanond, Pitiyanuwat & Archwamety, 2006). This perspective focuses more on the learning process in the mind of children. One learning theory that has been originated in light of cognitivism, is the theory of constructivism (Ackermann, 2001). This theory puts forth that knowledge is not simply passed on from teacher to student, but is formed by the students themselves when they interact with their environment (Tangdhanakanond, Pitiyanuwat & Archwamety, 2006; Sutherland, 1992). In other words, students are active learners who construct their own knowledge. They learn by doing and by interpreting their interactions with their environment, so they learn the most when they can actively perform their tasks. Since this method better fits the individual needs of students, their learning process in enhanced and deeper learning is established (Van der Meij, Kemps, Hoogland & Rutten, 2005).

Technology can be used to tailor education to the individual learning needs of pupils (Melhuish & Falloon, 2010). While with traditional teaching methods children have to work with the information given by their teacher and the prescribed literature, technology allows children to use information they collect themselves. Accordingly, tablets can be seen as a learning method that actively employs the constructivistic principles (Bellefeuille, 2006). Due to these tablets, children are able to look up information on their own and can hereby construct their own knowledge. Instead of copying behaviour and knowledge from their teachers, tablets encourage them to create and obtain their own knowledge. Also, tablets offer the possibility to provide students with information at all times (Melhuish & Falloon, 2010; Johnson, Levine, Smith & Stone, 2010). In addition, tablets facilitate other opportunities for children, such as manipulating and transforming texts to meet their needs (Eagleton & Dobler, 2007). This makes their experience with reading more engaging and interactive (Larson, 2010). Assuming that children learn more from an active way of learning than from a passive one, one would expect that tablet usage has a positive effect on the achievements of children.

Although empirical studies of these proposed mechanisms are non-existent, the general effect of tablet usage in schools on achievements has been studied. A meta-analysis of 110 studies that looked into the effect of mobile-integrated education on achievements shows that the effect size of learning achievements with tablets is 0.615 (Sung, Chang & Liu, 2016). This means that in terms of learning outcomes, tablets are significantly more effective than traditional methods that only utilize pen and paper. Another meta-analysis of both quantitative and qualitative studies reports a moderate to strong effect size for the effect of tablet usage on student achievements (Tamim, Pickup, Borokhovski, Bernard & El Saadi, 2015). Regarding reading outcomes specifically, an experiment with e-book reading exhibited significant progress in children's word comprehension and reading ability, as opposed to those who used printed texts (Korat, 2010). Lastly, a meta-analysis into reading outcomes shows that the electronic tools and applications in schools positively affect reading outcomes of students (Cheung & Slavin, 2012).

Social constructivism

The idea of constructivism can be extended by the theory of social constructivism. This theory argues that social participation is one of the main activities through which students learn and it opposes the notion that learning is an individual matter (Vosniadou, 2001; Lemke, 2001). Children obtain and create their knowledge through their interaction and engagement with others. Interaction ensures that children not only reflect on their own ideas, but also on those of their peers. Sharing their ideas and knowledge encourages them to critically reflect on information, hereby enhancing their learning processes and performance (Strommen & Lincoln, 1992; Smith, Wood, Adams, Wieman, Knight, Guild & Su, 2009).

Tablet usage enriches these interactions and discussions among peers, which in turn improves their performance. With traditional teaching methods, the knowledge children obtain is likely to be very similar, since they all get their information from similar sources (i.e. the same teacher or books). When children make use of tablets, their knowledge and ideas are more likely to be distinctive. Tablets offer children the opportunity to look up information on their own and hereby construct their own knowledge. This leads to more diversified ideas and opinions among children, because the information they obtain comes from a variety of sources. Sharing these different perspectives, instead of more alike perspectives, enriches and fosters the discussions and reflections among peers. That is, during these discussion students will encounter more opposing and diversified views. This requires higher-level thinking as students need to reflect on more different ways of looking at a certain topic. In turn, reflecting on these perspectives enhances their learning process because children obtain and create their knowledge primarily through their interaction with others.

Empirical support for this proposition is found by a qualitative study into iPads at schools (Flewitt, Messer & Kucirkova, 2015). The authors demonstrate that literacy activities with iPads increases motivation and concentration among children, enriches interaction and collaboration among peers, and lead to high levels of achievements. Another quantitative study examined the effect of the 'Write to Learn' method on literacy achievements (Genlott & Grönlund, 2016). Primary schoolers used ICT tools to write texts which they afterwards could discuss and improve together with classmates and teachers via digital formats. The central factor addressed here is how children can interact with peers and teachers, hereby providing social meaning and deeper learning. Two control groups are used; one group using the traditional method (no ICT) and the other using ICT individually without feedback. They find that their treatment group has the highest scores on literacy. Those that individually used ICT without feedback performed the worst, leading to the conclusion that access to technology alone is not enough. In other words, the use of ICT tools is more beneficial than traditional methods but only if they allow for interaction among peers and teachers.

Based on the theory of (social) constructivism, I expect that the use of tablets in schools exerts a positive effect on the reading achievements of children. In sum, my reasoning is three-folded a) it better fits the individual needs of students, b) it offers them greater opportunities to actively construct their own knowledge and c) it enriches the discussions and reflections among peers. Hence, hypothesis 1 is as follows:

H1: The use of tablets in primary schools has a positive effect on the reading achievements of children.

Gender differences

Research suggest that the ideas related to (social) constructivism might work differently for boys and girls. That is, how tablets fit the individual needs of students, how children react to an active way of learning, and to what extent discussions and reflections are enriched. Based on my theoretical considerations and prior empirical studies, I argue that tablet use is more profitable for boys than girls (shown graphically in figure 1). Below I discuss in detail why all three aspects are of greater advantage to boys than to girls when it comes to reading achievements.

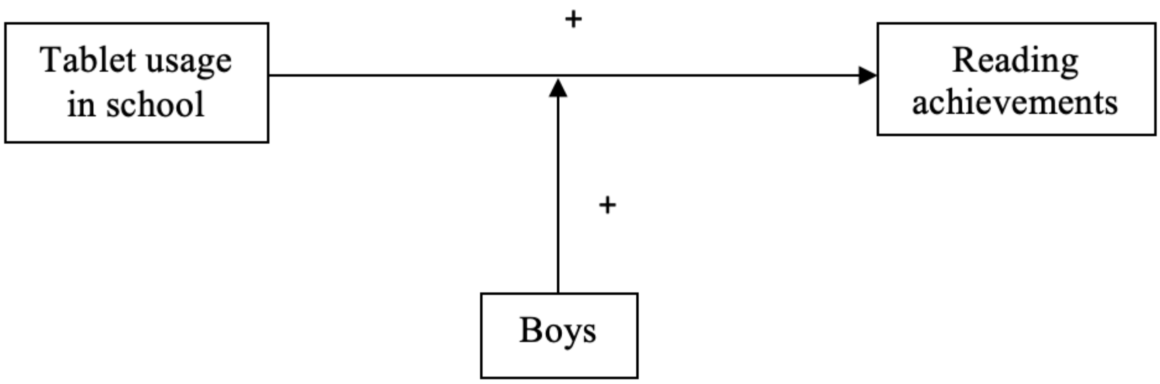

***Figure 1.* The moderating role of gender for the effect of tablet usage on reading achievements**

How *tablets better fit the individual needs* of students is more beneficial to low-achievers than to high-achievers. Tablets reduce cognitive task demands of assignments due its often repetitive and interactive nature (Outhwaite, Gulliford & Pitchford, 2017). This allows children to utilize different trial-and-error strategies and stimulates them to continue with hard assignments, which is found to improve achievements. Reduction of demands is something that turns out to be especially helpful for the performance of low-achievers, as compared to high-achievers (Outhwaite, Gulliford & Pitchford, 2017). When it comes to reading skills, boys can be considered low achievers since girls are often better in reading than boys (Lynn & Mikk, 2009). Most heard explanations for this discrepancy are the differences in reading frequency, interest in, enjoyment and attitudes towards (Torppa, Eklund, Sulkunen, Niemi & Ahonen, 2018). Girls more often like to read for pleasure than boys. If you are interested in reading, you have more opportunities to practice your reading skills (Gersheron & Holt, 2015). Less time invested in practicing reading then leads to fewer reading skills (Torppa, Eklund, Sulkunen, Niemi & Ahonen, 2018). For this reason, it often true that poor readers are those that have less interest in reading (Mol & Bus, 2011). Now, if boys are less interested in reading, they also have less experience with reading. That is why reading is more challenging to boys than it is to girls. Tablets then might be able to reduce these demands for reading due to its repetitive and interactive features. Given the finding that a reduction of demands is particularly helpful to low-achievers, reduced demands for reading due to tablet use will probably be most helpful to boys.

Studies into the benefits of individualized learning for boys and girls reading outcomes respectively are non-existent. What is available are a few studies into the effect of this for low- and high-achievers. Looking at this effect in more general terms, one study found that a more student-centered teaching strategy was most beneficial for low-achievers in terms of perceived learning (Nouri, 2016). This strategy better fitted the individual needs of students because it incorporated more interaction and allowed students to work at their own pace. Hereby, it reduced the demands for solving assignments, leading to an increase in perceived learning for especially the low-achievers. Another study regarding tablets only, found that the reduced cognitive demands due to tablet use were especially helpful for the mathematical achievements of low-achievers as compared to high-achievers (Outhwaite, Gulliford & Pitchford, 2017).

Another aspect of constructivism that might have a different impact is *how children respond to an active way of learning*. Research demonstrates that boys and girls have unique learning styles (Carrier, 2009). In traditional classroom environments students learn the most when they sit still, take notes, listen carefully and are able to multi-task (Carrier, 2009). Girls reported to fit better in these traditional classroom environments since these assets are more in line with their learning style. Natural assets that bring boys to learning are impulsivity, single-task focus and learning by doing (King & Gurian, 2006; Van de Gaer, Pustjens, Damme & Munter, 2007). Zooming in on ICT-based activities in particular, girls prefer to have an explanation given to them, while boys favour to explore this by themselves (Hornstra, van der Veen, Peetsma & Volman, 2015). In order to enhance the learning process of boys, it is suggested to provide more action-oriented activities (Taylor & Lorimer, 2003). Tablet usage can be characterized as an action-oriented activity, as it offers a range of applications and there are no fixed guidelines on how to complete a task when using the tablet (McPhee, Marks & Marks, 2013). For this reason, tablets better accommodate for the aforementioned typically male assets than for the female ones. The active way of learning and constructing your own knowledge instead of copying it from others, appears to be something that comes boys more naturally. Therefore, it is likely that tablet usage will foster the advancements in reading more for boys than for girls.

Empirical evidence for this proposition is found by the authors of the aforementioned 'Write to Learn' study (Genlott & Grönlund, 2016). They investigated if the Write to Learn method had a stronger effect for boys. They hypothesized that due to the more structured approach and technical tools supporting an active learning-oriented peer culture, the Write to Learn method would be more profitable for boys. The results indeed show that for literacy outcomes, boys gain more from the Write to Learn method than girls. Another study examined the effect of iPad usage in primary schools on children's engagement (McPhee, Marks & Marks, 2013). Despite the fact that this study investigates engagement instead of achievement, its theoretical considerations follow my ideas on how active ways of learning fit boys better. They found that engagement increased in classes were the iPad was used, and that this effect was considerably stronger for boys than girls. The authors suggest that tablets are potential tools to reduce the gender gap in achievements since increased engagement is likely to positively influence achievements.

At last, boys and girls may differ in *how they make use of the tablet* itself. This can best be explained by the stereotype threat paradigm (Steele, 1997; Steele & Aronson, 1995). People that belong to a group to which a negative stereotype is attached can suffer from performance impairments when their negative stereotype becomes salient. With regards to technology, the stereotype exists that boys are better with computers than girls (Koch, Müller & Sieverding, 2008; Smith, Morgan & White, 2005). Due to this stereotype boys and girls are treated differently; people encourage boys more than girls to use technological devices and are more confident that they know how to use it (Vekiri & Chronaki, 2008). It also influences the preferences and self-image of children. Boys gain greater interest in and become more confident with computers and tablets than girls. When it comes to computer or tablet related tasks in school, the negative stereotype becomes salient for girls and the positive stereotype for boys.

As a result, boys make better use of tablets than girls. They hereby create an advantage over girls in terms of knowledge, since the knowledge they obtain depends on how they make use of the tablet. In other words, the positive effects of tablet use on achievements will be greater for boys.

Numerous studies found evidence for this advantage of boys over girls with regards to technology usage, skills and interest. Research shows that while girls use ICT primarily for well-known tasks, boys use it for higher-level tasks (e.g. programming) (OECD, 2005). They show that the more advanced the task is, the greater the gap is. In addition, it is found that boys use computers and internet more often, have greater experience with it, consume more time online, have more interest in computer activities, are more positive towards it, and have greater motivation for obtaining digital skills (Broos, 2006; OECD, 2003; Sanders, 2005). How this advantage over technology then influences achievements is investigated in the Write to Learn study (Genlott & Grönlund, 2016). The authors found that the technical tool was most profitable for the literacy outcomes of boys. Another empirical study concludes that tablet computer technology in class can overcome the gender gap in engagement, as the engagement levels of boys increased significantly more than those of girls (McPhee, Marks & Marks, 2013). They suggest that higher levels of engagement will also lead to greater achievements.

Given these insights on how boys and girls profit differently from and respond to tablet usage, I expect a different effect of tablet use on school achievements for boys as compared to girls. Tablet usage will have a greater positive influence on the reading achievements of boys than those of girls. Hence, hypothesis 2 is as follows:

H2: The positive effect of use of tablets on reading achievements is greater for boys than for girls.

Methods

Data

In order to investigate my research questions, I use data from the Progress In International Reading Literacy Study (PIRLS, 2016). This is an international comparative quantitative study into the reading performance of students in grade four of primary education. The study has been conducted in 2001, 2006, 2011 and 2016 in over forty countries. For this research I utilize data from the Dutch subsample of 2016. The dataset contains rich information on students' reading achievements and the context of the different educational systems, school organizational approaches and instructional practices. Data is retrieved via reading tests and questionnaires regarding the school curriculum and background information of the students.

The study uses random samples to ensure representativeness of the national target population (Woessmann, 2016). Specifically, they utilize a two-stage cluster sampling design (LaRoche, Joncas & Foy, 2016). In the first stage, random samples are drawn of schools in each country. The sampling strategy took into account two stratification variables in the Netherlands; the socioeconomic status level of the school based on students' backgrounds and the degree of urbanization (i.e. how urbanized the area is in which the school is located). Within each sampled school, they randomly took one classroom per target grade (grade four, or group six in Dutch). Approximately 4000 students must take the test to obtain reliable results. The number of schools depends on the average number of students per class. The Dutch sample included 22 schools for the test measurement and 150 schools for the main measurement. It also included two reserve schools per selected school which corresponded as far as possible with regard to stratification variables of the originally selected school (Gubbels, Netten & Verhoeven, 2017).

In order to get a representative sample of the Dutch schools, at least 50% of the schools in the main sample need to participate in the survey. When the originally selected school refuses to participate in the study, the reserve school can be approached. The final response after approaching the reserve schools should be at least 85%. Only if both response requirements are met, the sample is representative. In total, 4206 students distributed over 226 classes in 132 Dutch schools participated in the PIRLS test. The final response rate was 88% and the response rate without reserve schools was 67%. Since this complies to both response requirements, the results are representative for Dutch fourth grade pupils (Gubbels, Netten & Verhoeven, 2017).

The research has been announced in magazines, digital newsletters and on various websites. Schools that were drawn from the sample received a letter with the request to participate in the research in the beginning of the 2015/2016 school year. Schools were then approached by phone to ask if they were willing to participate. Test leaders of the Dutch Expertise Center visited the schools in the period from 14 March, 2016 to 27 May, 2016 on a date chosen by the school. On this date, all fourth-grade students of that school took the PIRLS test and filled in the student questionnaire. The test leader was responsible for the documentation of absenteeism of students. Reasons for absence were that students had left the school permanently, that parents objected to participation, or that students could not participate due to a physical disability, special educational needs or insufficient knowledge of the Dutch language. During the tests, teachers completed the teacher questionnaire which was either given to test leader or send afterwards. The school leader questionnaire was completed by the director or supervisor of the school and the parent questionnaire by a parent or guardian of the pupil, which both were send afterwards (Gubbels, Netten & Verhoeven, 2017).

From the 4206 students that participated, 2989 are used in this analysis. Since I need data about the tablet use of students for this research, 115 cases of the original 4206 drop out because data on their tablet use is non-existent. Of the 4091 that remain, 62 extra cases are lost due to missing values on the students' socioeconomic background variables. Finally, there are another 569 missing values for the financial situation of the school, and 471 for the SES composition of the school. This results in 2989 valid cases that I will use for my analysis.

Measurement instruments

To create a measurement instrument for the dependent variable *reading achievements*, the plausible value scores of the overall reading skills are being used. Plausible values are the scores that result from the PIRLS test. The PIRLS test consists of twelve texts that are divided via a test rotation system into sixteen test books, each containing two texts and related questions. Students are randomly assigned to a test booklet. Because each student only answers the questions for two texts and not all texts are equally difficult, no individual learning scores are calculated. To enable comparison between students and countries, PIRLS uses Item Response Theory. This means that a skill distribution is made for each student based on the skill of the student and the difficulty of an item. Five estimated values, plausible values, are taken from this distribution. These values form a student's reading skills score. This score ranges from approximately 300 to 700 and is standardized on a scale of 500. Scores below 400 are low; between 470 – 480 are intermediate; between 545 – 555 are high; and above 620 are advanced. The dataset contains plausible values of four different aspects of reading and of their overall reading score (i.e. the four different aspects combined). Since I am interested in reading skills in general, I will utilize the plausible values for the overall reading score only. For the sake of simplicity, I use the mean of these five single plausible value scores. The variable reading achievements is thus the overall mean of these single mean scores. This variable does not contain any missing values.

Student level predictors

The independent variable *tablet use* is measured by the question how often respondents use a computer or tablet in school for schoolwork (including classroom tasks, homework, or studying outside of class). Participants could indicate that they 'never or almost never', 'once or twice a month', 'once or twice a week', or 'every day or almost every day' use tablets in school for these purposes. Higher values represent greater tablet use. Although this variable does not distinguish between tablets or computers, this is not too problematic. The mechanisms that underly (social) constructivism are likely to work similarly for computers. Computers offer children the same opportunities to construct their own knowledge, and hereby also enrich discussions among peers. This suggests that the effect of tablets on achievements will be the same for computers. This variable contained 87 omitted or invalid values, and 28 system missing values. Omitted or invalid values are incorrect values that have been set to missing by the user, and system missing values are non-existent values that have been set to missing by SPSS itself. These 115 cases in total have been set as missing values and have thus been excluded from my analysis.

The moderator used in this study is the *gender* of the students. It is coded so that boys have a 1 and girls have a 0, thus girls are the reference category in this study. This variable did not have any missing values.

Student level control variables used in this study are *age* and the *socioeconomic background of the student*. The age of the participants is their age at the time the survey was taken. There were no missing values for this variable. Although there is only slightly variance in the age of the respondents (ranging from 8.5 to 12.5), this could be influencing the outcomes. Older students are found to perform better in school achievements (Schneeweis & Zweimüller, 2014). Moreover, older children might have greater experience with tablets use and therefore make better use of it. In turn, this would positively influence their reading achievements.

Students social economic background is controlled for since students from lower social economic backgrounds tend to perform less than children from higher social economic backgrounds (e.g. Saifi & Mehmood, 2011). Moreover, those from lower socioeconomic backgrounds might profit more from tablet use since they can be considered the low-achievers (Outhwaite, Gulliford & Pitchford, 2017).

Students social economic background was not asked directly in the questionnaire. On the contrary, in the parent's questionnaire, parents were asked for their educational and occupational status. However, these variables both contained a high number of missing values. Including one of these variables would mean that I would lose half of my cases. Hence, I use two alternative indicators to derive the students' their socioeconomic background. Socioeconomic status consists of different dimensions, most prominently an economic and cultural dimension (Bloom, 1964; Keeves, 1972). Possessions at home are indicators for the economic dimension (Yang, 2003). For the cultural dimension, the number of books at home is often used as indicator (Yang, 2003; Sieben & Lechner, 2019; Lubienski & Crane, 2010).

Following these arguments, I use the economic and cultural resources available at home as indicators of students' their socioeconomic background.

For the economic dimension, I utilize the number of possessions at home. Students were asked 'Do you have any of these things at your home?'. They could answer either yes or no for 'computer or tablet', 'study desk', 'own room', and 'internet connection'. I constructed one variable in which the number of possessions is reflected (i.e. the number of times they ticked yes). There were some missing values on the original variables, ranging from 37 to 51 missing values. These were either omitted or invalid, or system missing values. Missing values on the original variables were interpreted as participants not having that specific possession at home. This resulted in 67 missing values on the final variable, which were listwise deleted prior to the analysis. The cultural dimension was measured by the question 'About how many books are there in your home? (Do not count magazines, newspapers, or your schoolbooks)'. Answers categories were on a 5-point scale: 'none or very few (0 – 10 books)', 'enough to fill one shelf (11 – 25 books)', 'enough to fill one bookcase (26 – 100 books)', enough to fill two bookcases (101 – 200 books)', 'enough to fill three or more bookcases (more than 200)'. In my analysis this variable is treated as a continuous variable. It contained 30 omitted or invalid values, and 28 system missing values. This means that in total 58 cases have been listwise deleted before the analysis.

<u>School level predictors</u>

The first school level control variable that has been incorporated is *shortages in schools*. It is important to take this into account because it could be a confounding variable. Shortages in schools are likely to affect the reading achievements of children, as well as the tablet use in that school. Schools that are confronted with shortages cannot invest in tablets for educational purposes, and thus do not use tablets. These shortages are also likely to affect the achievements of students; a lack of resources such as (qualified) teachers or instructional materials, negatively affects students' performance (Scholten & Wolbers, 2019). If not controlled for this situation, the results might yield an illusory positive relationship between tablet use and achievements. I constructed a scale from 13 variables on school shortages. School principals were asked 'How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following?'. The following consists of shortages on general resources, being instructional materials, supplies, school buildings, heating systems, instructional space, technological staff, audio-visual resources, computer technology, and resources for students with disabilities. Further, it consists of shortages on resources specifically for reading, being teachers specialized in reading, computer software for reading, library resources and instructional materials. For each indicator they can choose ‘not at all’, ‘a little’, ‘some’, or ‘a lot’. Higher values thus represent greater shortages. For each question data was missing ranging from 524 missing values to 602 missing values. I ran one confirmatory factor analysis with all thirteen items of which I extracted one factor (see appendix A1). Hereafter, I ran another factor analysis with the eight items of the first factor analysis that loaded above .4 (see appendix A2). The second factor analysis resulted in a reliable scale with an alpha of .777 (see appendix B1). Hence, I utilize the second factor score as my shortage variable. Values on the original variables were treated as missing for the construction of the scale. The final scale has 597 missing values, which were missing listwise deleted prior to the analysis.

The second school-level variable is the *school socioeconomic composition*. This variable is included for similar reasons as to why the students’ SES background is included. The SES of the school has a positive impact on educational performance, even when you control for family SES (McConney & Perry, 2010). School principals are asked to indicate their schools’ composition by student background. They were asked: ‘Approximately what percentage of students in your school have the following backgrounds?’. Two backgrounds were distinguished: a) Come from economically disadvantaged homes, b) Come from economically affluent homes. For each category they could answer ‘0 to 10%’, ’11 to 25%’, ’26 to 50%’, or ‘More than 50%’. Given the data on these two variables, researchers from the PIRLS study have constructed one school composition variable, which I use as my school SES variable. Herein, schools are rated as ‘more disadvantaged’, ‘neither more affluent nor more disadvantaged’, or ‘more affluent’. Again, higher values thus represent a higher SES composition. The variable SES composition of the school has 1066 missing values, which were missing listwise deleted.

Description of measurement instruments

Descriptions of the measurement instruments can be found in table 1. Reading skills are based on the mean of the plausible value scores. This score ranges from 356.05 to 737.13, with the average score approximately in the middle (mean = 549.25, SD. = 54.49). In comparison with other participating countries this average score can be considered a middle to high score. South Africa scores lowest on average with a score of 320, and the Russian Federation scores highest with a score of 581 (PIRLS, 2018). The majority of the students uses tablets once or twice a week at school (mean = 2.43, SD = 1.11). Given the fact that tablet use is measured on a 4-point scale, the SD can be considered relatively large. This indicates that the amount of tablet usage differs greatly between schools. Gender is distributed equally in this sample, with 50% of the students being male. Their age ranges from 8.5 to 12.5 years old, with a mean of 10 years. For students' their socioeconomic background, the economic indicator reflects their number of possessions at home. Most students have all four possessions, meaning they have a computer or tablet, study desk, own room and internet connection (mean = 3.83, SD. = 0.45). The cultural dimension of students' socioeconomic background is measured by the number of books in their home, which for most students is one bookcase full (26 – 100 books). This is approximately in the middle of the 5-point scale, suggesting that within the cultural dimension no particular background status is overrepresented (min. = 1, max. = 5, mean = 2.93, SD. = 1.08). Further, most school leaders indicate that they do not have shortages in their schools (min. = -1.06, max. = 2.71, mean = 0.06, SD. = 0.96). At last, the majority of the school principals indicate that their school is composed of students from more affluent backgrounds (min. = 1, max. = 3, mean = 2.47, SD. = 0.65).

Table 1

Description of measurement instruments

	N	Min.	Max.	Mean	SD.
Reading score	2989	356.05	737.13	549.25	54.78
Tablet use	2989	1	4	2.43	1.11
Male	2989	0	1	0.49	n.a.
Age	2989	8.5	12.5	10.04	0.46
Economic SES	2989	0	4	3.83	0.45
Cultural SES	2989	1	5	2.93	1.08
Shortage scale	2989	-1.06	2.71	0.06	0.96
SES school	2989	1	3	2.47	0.65

Note: Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.

Analytical strategy

A two-level multilevel model has been developed to examine the influence of tablet use on reading achievements. In many educational studies such as the PIRLS data, a hierarchical structure exists (Sun, Bradley & Akers, 2012). For instance, students are nested within classes, and classes are nested within schools, and schools are nested within regions, and so forth (Hox, 2002). Ordinary least-squares multiple regression cannot be applied, since the assumption of independent observations is hereby violated (Luke, 2004). OLS regression has a downward bias in standard errors. As a result, even limited non-independence can lead to an increase in type 1 error probability (Osborne, 2017). This means that you reject your null hypothesis while actually this hypothesis is true. Multilevel modelling is most appropriate for such hierarchical data because it accounts for the non-independence of observations (Sun, Bradley & Akers, 2012). That is, for each level in the hierarchical structure a sub model is created. These sub models display the relationship between variables at a certain level, and point out how variables at higher levels affect features and tendencies at a lower or corresponding level.

The multilevel model has been conducted in four steps, similarly to the recommended strategy for Hierarchical Linear Modelling (Hofmann, 1997; Raudenbush & Bryk, 2002). Parameters were estimated based on Maximum Likelihood Estimation with robust standard errors (Koth, Bradshaw & Leaf, 2008). First, I ran a random intercept and slopes model without including any predictors (the null model) (Hox, 2010). This model allowed for the intercepts to randomly vary between schools. Hereby, information was provided of the variances within and between schools for reading achievements. The null model has been estimated based on the equations below.

Level 1 equation:

$$\gamma ij = \beta 0j + eij$$

Level 2 equation:

$$\beta 0j = \gamma 00 + \mu 0j$$

Reading score γij represents the dependent variable, which is the reading achievement for student *i* in school *j*. $\beta 0j$ is the intercept that represents the average reading achievements of the school j, this intercept varies at the school level. eij is the error term that represents a unique effect associated with student *i* in school *j*. It is assumed that this error term is normally distributed and has a mean of zero. In the level 2 equation, $\gamma 00$ is the intercept that depicts the mean of reading achievement and $\mu 0j$ is the (random) residual error at the class level. Again, it is assumed that this residual error has a mean of zero and is independent from the residual errors at the student level.

In addition to testing the variance components for statistical significance, the intraclass correlation coefficient (ICC) has been calculated for the null model. This allowed for further evaluation of the level of non-independence in the outcome. The ICC is defined as the expected correlation between "any two randomly chosen students in the same school" and is computed as the proportion of variation in the level 1 outcome "explained by the grouping structure" (Heck, Thomas & Tabata, 2014, p.8). The ICC is calculated by the formula below, in which σ_e^2

refers to the variance of the lowest-level errors *eij*, and σ_u^2 is the variance of the highest-level errors *u0j* (Hox, 2010):

$$\rho = \frac{\sigma_{u0j}^2}{\sigma_{u0j}^2 + \sigma_{eij}^2}$$

Hereafter, I added my level one predictors (i.e. student level) to the model. These student-level predictors are treated as fixed and assumed to have the same impact across schools. This model illustrates the effect of tablet use on reading achievements. The second model is estimated is based on the equations below.

Level 1 equation:

$$\gamma ij = \beta 0j + B1j\ (tabletuse) + \beta 2j\ (male)ij + \beta 3j\ (age)ij + \beta 4j\ (ses\ eco)ij + \beta 5j\ (ses\ cul) + eij$$

Level 2 equation:

$$\beta 0j = \gamma 00 + \mu 0j$$
$$\beta 1j = \gamma 10$$
$$\beta 2j = \gamma 20$$
$$\beta 3j = \gamma 30$$
$$\beta 4j = \gamma 40$$
$$\beta 5j = \gamma 50$$

B1j is the coefficient of tablet use, which measures the relationship between tablet usage in school and student reading achievement in school *j* (and so on for the other coefficients). $\gamma 00$ is the intercept that depicts the mean of reading achievement. In this model, student-level predictors are treated as fixed and assumed to have the same impact across schools, therefore, $\beta 1j = \gamma 10, \beta 2j = \gamma 20, \beta 3j = \gamma 30, \beta 4j = \gamma 40, \beta 5j = \gamma 50.$

Then my level two predictors (i.e. school level) were added to the model. These school-level variables are the random effects. In model 3 the compositional variables as predictors of the variation in intercepts are incorporated. Fitting the multilevel model at both students and school levels leads to a more meaningful interpretation of the variables (Ma, Ma & Bradley, 2008). The model is based on the equations that are presented below.

Level 1 equation:

$$\gamma ij = \beta 0j + B1j\ (tabletuse) + \beta 2j\ (male)ij + \beta 3j\ (age)ij + \beta 4j\ (ses\ eco)ij + \beta 5j\ (ses\ cul) + eij$$

Level 2 equation:

$$\beta 0j = \gamma 00 + + \gamma 01(shortage)j + \gamma 02(ses\ school)j + \mu 0j$$

$$\beta 1j = \gamma 10$$

$$\beta 2j = \gamma 20$$

Interpretation of the coefficients is similar to those in the second model. $\gamma 01$ is the coefficient of shortages in a school, measuring the relationship between shortages and school reading achievement in level 2. $\gamma 02$ is the coefficient of the socioeconomic status of the school, which measures the relationship between the schools' socioeconomic status and school reading achievement. Again, student-level predictors are treated as fixed, hence $\beta 1j = \gamma 10$, $\beta 2j = \gamma 20$, $\beta 3j = \gamma 30$, $\beta 4j = \gamma 40$, $\beta 5j = \gamma 50$.

At last, I ran another model with now the interaction term for gender included as well. This full model was created to investigate whether the effect of tablet use on achievements is different for boys and girls. The following equations were used for this model:

Level 1 equation:

$$\gamma ij = \beta 0j + B1j\ (tabletuse) + \beta 2j\ (male)ij + \beta 3j\ (age)ij + \beta 4j\ (ses\ eco)ij + \beta 5j\ (ses\ cul) + \beta 6j\ (tablet * male)ij + eij$$

Level two equation:

$$\beta 0j = \gamma 00 + + \gamma 01(shortage)j + \gamma 02(ses\ school)j + \mu 0j$$

$$\beta 1j = \gamma 10$$

$$\beta 2j = \gamma 20$$

The four models have been compared to each other via goodness-of-fit indices. These indices indicate how well the model fits the data compared to the other models. For each model, the -2 Log Likelihood (-2LL) and the Akaike's information criterion (AIC) are shown. -2LL is a chi-square test, so for each model the χ^2 is given. Hereafter, I calculated the change in χ^2 and the change in the degrees of freedom with:

$$\chi^2_{change} = (-2LL_s) - (-2LL_1)$$
$$DF_{change} = df_s - df_1$$

Given the DF_{change} one can estimate whether the χ^2_{change} is statistically significant by looking at the chi-square distribution table (see appendix C) (Gravetter & Wallnau, 2016). If the χ^2_{change} is greater than this number in chi-squared distribution, then the bigger model (1) fits significantly better than the smaller model (s). The smaller model is nested in the bigger model in terms of the number of parameters. AIC values can be compared more easily; lower AIC values simply suggest better models.

Additionally, the effect of adding the student-level explanatory variables to model has been assessed by means of the explained variance R^2. Since SPSS does not report the explained variance with MLM, the proportion of variance explained at the student level has been estimated with the formula below (Raudenbush & Bryk, 2002).

$$R_1^2 = \left(\frac{\sigma^2_{e|b} - \sigma^2_{e|m}}{\sigma^2_{e|b}}\right)$$

$\sigma^2_{e|b}$ refers to the lowest-level of residual variance for the baseline model (i.e. the intercept-only model), and $\sigma^2_{e|m}$ refers to the lowest-level residual variance for the comparison model (Hox, 2010). The proportion of variance explained at the class level has been calculated with:

$$R_2^2 = \left(\frac{\sigma^2_{u0|b} - \sigma^2_{u0|m}}{\sigma^2_{u0|b}}\right)$$

Herein, $\sigma^2_{u0|b}$ refers to the second-level residual variance for intercept-only model, and $\sigma^2_{u0|m}$ to the second-level residual variance for the comparison model.

Prior to performing the analysis, model assumptions were precisely checked. A plot of the residuals versus the predicted scores showed that the assumption of a linear relationship and homogeneity of the residuals was not violated (Luke, 2004; Raudenbush & Bryk, 2002). Inspection of the normal quantile plot of the residuals yielded that the prediction errors given tablet use were normally distributed.

Results

This study utilized the Dutch sample of the PIRLS study (2016) to explore the factors impacting reading achievements from a Multilevel Modelling Approach. Four different models have been estimated. Table 2 presents the coefficients and standard errors for the four different models. The results of the analysis for each model are discussed in this section.

The null model

The null model allows the intercepts to randomly vary between schools. Since no predictors are incorporated in the model at level 1, the intercepts are equal to the school means for the level 1 outcome variable (i.e. reading achievements). As can be seen in table 2, the fixed effect of reading achievements is 545.069. Random effects of the null model are also presented. The within-schools variance in reading achievements is σ^2_{eij} = 2758.479. The between-schools variance, which reflects the variation in intercepts, is σ^2_{u0j} = 402.671. To determine whether there is significant variation to be explained at Levels 1 and 2, a Wald Z-test is conducted. Based on the findings of the Wald Z-test in table 3, both variance components are statistically significant ($p<.001$).

The ICC for this model .127 (see Appendix D for calculations). ICC values above .05 are often considered an indicator of a non-trivial amount of non-independence. The ICC of .127 for the current model thus suggest that MLM is required.

Table 2

Fixed effects, random effects and model fit indices for each model

	Model							
	Null model		Second model		Third model		Final model	
Parameters	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*	Coefficient	SE
Fixed effects								
Intercept	545.069***	(1.975)	684.779***	(19.895)	643.941***	(23.182)	643.339***	(23.357)
Tablet usage			-.903	(.772)	-.419	(.876)	-.243	(1.205)
Male			-9.685***	(1.588)	-9.097***	(1.811)	-8.257	(4.348)
Age			-18.992***	(1.802)	-17.562***	(2.009)	-17.545***	(2.010)
SES eco			6.035**	(1.817)	5.760**	(2.065)	5.759**	(2.064)
SES cul			12.474***	(.765)	11.785***	(.880)	11.789***	(.881)
Shortages					-2.156	(1.895)	-2.164	(1.895)
SES school					11.942***	(2.612)	11.944***	(2.612)
*Tablet * male*							-.346	(1.629)
Random effects								
Level 1 error variance	2758.479***	(61.120)	2465.970***	(55.905)	2378.379***	(62.650)	2378.325***	(62.649)

Level 2 error variance	402.671***	(63.464)	267.909***	(46.516)	217.089***	(46.097)	217.158***	(46.108)
Model fit								
ICC [a]	.127							
-2LL [b]	45473.379		43083.512		31483.940		31843.895	
AIC [c]	45479.379		43099.512		31863.940		31865.895	
R_1^2 [d]			.106		.138		.138	
R_2^2 [e]			.333		.461		.461	

Note: *p < 0.05; **p < 0.01; ***p < 0.001 (two-tailed test). Standard errors are between parentheses. Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.
[a] ICC = intraclass coefficient
[b] -2LL = Log Likelihood
[c] AIC = Akaike information criterion
[d] R_1^2 = Explained variance at the student level
[e] R_2^2 = Explained variance at the school level

The second model

Looking at the estimate of fixed effects for the second model in table 2, one can see that tablet use does not significantly predict reading achievements (B = -.903, p = .242). Gender is a significant positive predictor of achievements (B = -9.685, p<.001), indicating that within groups males tend to score lower on reading achievements than females. Age negatively influences reading achievements (B = -18.992, p<.001), which suggests that within schools, older students perform worse in reading than younger students. Also, both socioeconomic indicators are significant predictors of reading achievements within schools. Students from higher economic backgrounds tend to perform better on reading (B = 6.035, p<.01), and the same is true for students from higher socio-cultural backgrounds (B = 12.474, p<.001). Interpretation of the coefficients can be explained in this way: for every unit that age increases, the student reading score decreases with 18.99 points when holding all other variables constant ($\gamma 30$ = 18.99). Further, both the within- and between-schools variance component is statistically significant based on the results of the Wald Z-test (p<.001).

In order to estimate whether this model fits the data better than the null model, the fit indices can be utilized (for calculations of the fit indices see Appendix D). Based on a comparison of the -2LL and the AIC it can be concluded that this model fits better than the null model. The explained variance at the student level is .106. Thus, 10,6% of the variance in students reading achievements is explained by the predictors included in this model. The explained variance at the school level is .333, which means that 33,3% of the variance at the class level can be explained by the tablet use, gender, age and the socioeconomic background of the student.

The third model

For model three, the effects of the student-level variables tablet use, gender, age and socioeconomic background on students' reading achievements are similar to those in the second model. At the school-level, shortages in schools do not predict reading achievements in each school (B = -.419, p = .663). In contrast, the SES composition of a school is an effective predictor of the average reading score at each school (B = 11.942, p<.001). The interpretation of the intercepts of school level variables is as follows: for every unit increase in the SES composition of the school, student reading achievement increases with 11.94 points keeping all other variables constant ($\gamma 01$ = 11.942). Once again, the within- and between schools' variability remains significant given the Wald Z test (p<.001).

Fit indices are now utilized to estimate whether the third model is a better fit to the data than the second model (for calculations of the fit indices see Appendix D). Inspection of the -2LL and the AIC leads to the conclusion that the third model is a better fit than the second model. Assessing the effect of adding the explanatory variables to model results from the explained variance *R2*. This specific model explains 13,8% of the variance at the student level. The proportion of variance explained at the second level is 0.461. Thus, 46,1% of the variance at the school level is explained by the included predictors in the model. In comparison with the 33,3% of the second model, this shows that the greatest predictive power stems from the third model.

The full model

Findings of the full model in table 2 show that the interaction term is insignificant (B = -.346, p = .832), indicating that the effect of tablet usage on reading achievements does not work differently for boys and girls. Other fixed effects on both student- and school level remain similar to those in model three, as well as the significance of the variance components presented in table 9. There is however one exception, which is gender. Gender becomes insignificant with a p-value of .058.

Again, fit indices are being utilized (for calculations of the fit indices see Appendix D). Given the -2LL and the AIC of this full model it can be concluded that the third model is a better model than the full model. The explained variance at the student level remains similar to the third model with a value of .138. Likewise, the proportion of variance explained at the school level for the full model is also 46,1%. This similarity means that this model does not add any predictive power to the third model. Since the most parsimonious model (i.e. a model with the greatest predictive power with as few predictors as possible) is desired, model three can be considered a better model.

The insignificance of the interaction term and the model fit indices indicate that model three is the preferred model. Conclusions regarding my hypotheses are therefore based on model three. Given the findings of model three, my first hypothesis is rejected. That is, I find no evidence that tablet usage in schools has a positive effect on the reading achievements of children. Further, on the student level boys tend to perform worse than girls on reading, but the effect of tablet use on reading achievements does not work differently for boys and girls. Hereby, my second hypothesis is also rejected. I do find evidence for most of my control variables. On the student level I find that students' from higher socioeconomic backgrounds tend to perform higher than those from lower socioeconomic backgrounds. Contrary to my expectations, reading achievements decrease with increasing age.[1] On the school level, the SES composition of the school has a positive effect on the reading achievements of children. Shortages in schools do not predict reading achievements at the school level.

Conclusion and discussion

Only little is known about how tablets influence reading achievements. The purpose of this study therefore was to examine the effect of tablet usage in primary schools on the reading skills of students. I utilized the Dutch subsample of the PIRLS 2016 dataset, which contained extensive data on both tablet usage and reading achievements in primary schools. Drawing on theories of (social) constructivism and literature on tablet use in schools, I formulated two hypotheses. Contrary to my first hypothesis, tablet usage in primary schools does not have a positive effect on the reading achievements of children. The findings yielded a neutral relationship between tablet use and reading achievements. My second hypothesis is also inconsistent with my findings. I did not find that the effect of tablet use on reading achievements differs between boys and girls.

In contrast to my results, most prior studies found a positive relationship between tablet use and reading achievements (e.g. Sung, Chang & Liu, 2016; Korat, 2010). From a methodological perspective the discrepancy could be explained by the fact that my tablet variable also included computers. Given my theoretical considerations on constructivism, it seemed plausible that computers can have similar effects as tablets. However, tablets also share

[1] This negative effect could be explained from a theoretical perspective, as it is likely that older students more often form a selective group. For instance, older students are often students that have to repeat their school year, who possibly already have less reading skills.

certain advantages over computers that might be of importance here. Due to the smaller size of tablets, their mobile nature, and the rise of more social application software on it, tablets create more opportunities for collaboration than traditional PCs do (Henderson & Yeow, 2012; Haßler, Major & Hennessy, 2016). As said, collaboration enhances discussions and reflections among peers, which in turn is essential for learning. If computers allow less for these collaborations, it might be that this has led to an underestimation of the true effect. Arguing from a theoretical perspective, the difference in results might have to do with the technological skills of teachers during the time the survey was taken. Research has shown that technological competent teachers are needed in order for technological devices to be beneficial to children (Comi, Argentin, Gui, Origo & Pagani, 2017). Due to the novelty of tablets, teachers could lack relevant experiences and skills with tablets, or be unable to integrate it into their teaching process. This might be especially true for this research since the data is from 2016. Around this year the implementation of tablets into the Dutch educational system was relatively new. The absence of an effect could thus have been the result of teachers lacking relevant experience with tablets at that time, due to which they could not incorporate tablets correctly into the classroom.

The moderating role of gender in the context of technological devices has only been investigated by two previous studies. One study found that boys gained more from the implementation of an ICT method than girls (Genlott & Grönlund, 2016). Another research found that iPad usage increased engagement considerably more for boys than girls (McPhee, Marks & Marks, 2013). My results are not in line with these studies, as my interaction effect was insignificant. This suggests that there are no differences in how tablets influence achievements between boys and girls. However, the absence of an effect could have been the result of high levels of variation in how tablets are being used between both students and schools. For instance, are books completely replaced by tablets, or are tablets only used as extra support tool for students with reading difficulties (e.g. dyslexia)? Or on the contrary, are they used as extra material for students that need more challenging assignments? Also, the ratio of students to tablets could be of importance. If there are only a limited number of tablets available, then teachers have to decide how to divide and work with these tablets. These are just a few aspects that could vary between both students and schools. Deviations in how tablets are used individually, and how these are implemented into the school curriculum, could therefore have biased my results. The advantage for boys over girls in the use of tablets could thus still exist, but might not be exposed in this research due to high levels of variation. Accordingly, it cannot be said whether tablets are potential tools for reducing the gender gap in reading achievements.

The results of this study are not supportive of the theories of constructivism and social constructivism. Based on these theories I argued that tablets better fit the individual needs of students and offer them greater opportunities to actively construct their own knowledge, which in turn would enrich discussion among peers. As a result, reading achievements would be enhanced. I further discussed how these aspects of (social) constructivism would be more prone for boys, so that tablets would be more beneficial to boys than girls. Nevertheless, the lack of empirical evidence does not directly imply that tablets cannot actively employ these constructivistic principles. Before final conclusions can be drawn on these theories, strengths and limitations of this study need to be addressed.

One of the main strengths of this study is the adoption of a multilevel model approach. The nested structure of the data set violated the assumption of independent observations required for OLS regression. Due to the downward bias in standard errors, OLS regression would have led to an increase in type 1 error probability (Osborne, 2017). Multilevel modelling was the most appropriate method for this data because it accounted for the non-independence of observations. Also, the data set that has been used is highly relevant. It encompassed extensive data on both reading achievements and tablet use of Dutch primary schoolers. Hereby, it formed a large and representative sample of my target population. Finally, this was one of the first studies that looked at tablet use in general instead of merely reading from an e-book, and one of the first that used the Dutch primary school context (McPhee, Marks & Marks, 2013).

There are also some limitations that need to be acknowledged. As aforementioned, the fact that the tablet variable did not distinguish between computers and tablets might have biased the results. It is likely that computers allowed less for collaboration than tablets, hence, the true effect might have been underestimated depending on how often children meant computers instead of tablets. If a division was made between these two, I would have expected a positive effect of tablets on achievements. Therefore, I recommend future studies to utilize data exclusively on tablets, instead of data taken together with computers or other technologies. Another drawback of this study has to do with the lack of information on how tablets are being used individually and within schools. As said, high levels of variation could have biased my results. It might be that there is no general effect of tablet use on achievements, but that this effect depends on how tablets are being used individually or between schools as a whole. It would therefore be relevant for future research to utilize or collect data that contains more detailed information on these variations (i.e. what happens within schools and classes). Future studies could also reformulate my research question, by including tablet usage as random effect instead of a fixed effect. This means that you longer assume that tablets have the same effect across schools, but that this varies between schools. If such variations between students and schools become known, one can find out in which specific contexts tablets are beneficial for reading achievements.

The discussion of the limitations of this study and the scarce literature available on this topic illustrates that there is still abundant reason to study the influence of tablets on performance. More studies with improved research design are needed in order to provide a definite answer on whether the theories of (social) constructivism are applicable to the effect of tablet use on achievements. Findings of future studies can guide primary schools in the implementation of tablets into the school curriculum, so that the returns on tablets become most profitable to students. If tablets indeed turn out to be more beneficial to boys than girls, this could also help primary schools to address the current crisis in underachievement of boys in reading. That is, tablets could then be utilized to close the gap in reading achievements.

Literature

Ackermann, E. (2001). Piaget's constructivism, Papert's constructionism: What's the difference. *Future of learning group publication*, *5*(3), 438.

Beblavý, M., Baiocco, S., Kilhoffer, Z., Akgüç, M., & Jacquot, M. (2019). Index of Readiness for Digital Lifelong Learning: Changing how Europeans upgrade their skills. *Centre for European Policy Studies,* (No. 25419).

Bellefeuille, G. L. (2006). Rethinking reflective practice education in social work education: A blended constructivist and objectivist instructional design strategy for a webbased child welfare practice course. *Journal of Social Work Education*, 42 (1), 85-103.

Bloom, B.S. (1964). Stability and Change in Human Characteristics. New York: Wiley.

Broos, A., & Roe, K. (2006). The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents. *Poetics*, *34*(4-5), 306-317. doi:10.1016/j.poetic.2006.05.002.

Burden, K., Hopkins, P., Male, T., Martin, S., Trala, C. (2012). *iPad Scotland Evaluation.* Hull: University of Hull.

Carrier, S. J. (2009). Environmental education in the schoolyard: Learning styles and gender. *The Journal of Environmental Education*, *40*(3), 2-12. https://doi. org/10.3200/JOEE.40.3.2-12.

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review, 7*(3), 198-215. https://doi.org/10.1016/j.edurev.2012.05.002.

Churchill, D., Fox, R. M. K., & King, M. (2012). Study of affordances of iPads and teacher's private theories. *International Journal of Information and Education Technology*.

Cohen, M., Hadley, M., & Frank, M. (2011). Young children, apps & iPad. *US Department of Education Ready to Learn Program*, *200*, 5-10.

Comi, S. L., Argentin, G., Gui, M., Origo, F., & Pagani, L. (2017). Is it the way they use it? Teachers, ICT and student achievement. *Economics of Education Review*, *56*, 24-39. http://dx.doi.org/10.1016/j.econedurev.2016.11.007.

Gasparini, A. A., & Culén, A. L. (2012). Tablet PCs–An assistive technology for students with reading difficulties. In *ACHI 2012: The Fifth International Conference on Advances in Computer-Human Interactions* (pp. 28-34).

Gershenson, S., & Holt, S. B. (2015). Gender gaps in high school students' homework time. *Educational Researcher, 44*(8), 432-441. doi: 10.3102/0013189X15616123.

Druin, A., & Solomon, C. (1996). *Designing Multimedia Environments for Children: Computers, Creativity, and Kids*. Wiley Computer Publishing, John Wiley and Sons, Inc., One Wiley Drive, Somerset, NJ 08875.

Dundar, H., & Akcayir, M. (2012). Tablet vs. Paper: The Effect on Learners' Reading Performance. *International Electronic Journal of Elementary Education*, 4(3), 441- 450.

Eagleton, M.B., & Dobler, E. (2007). *Reading the web: Strategies for Internet inquiry*. New York: Guilford.

Faber, J. M., & Visscher, A. J. (2016). De effecten van Snappet. *Effecten van een adaptief onderwijsplatform op leerresultaten en motivatie van leerlingen*.

Flewitt, R., Messer, D., & Kucirkova, N. (2015). New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy*, 15(3), 289-310. http://dx.doi.org/doi:10.1177/1468798414533560.

Fokides, E., & Atsikpasi, P. (2017). Tablets in education. Results from the initiative ETiE, for teaching plants to primary school students. *Education and Information Technologies*, *22*(5), 2545-2563. doi: 10.1007/s10639-016-9560-3.

Furió, D., González-Gancedo, S., Juan, M., Seguí, I., Costa, M., & colleagues (2013). The effects of the size and weight of a mobile device on an educational game. *Computers & Education*, 64, 24–41. doi:10.1016/j.compedu.2012.12.015.

Genlott, A. A., & Grönlund, Å. (2016). Closing the gaps–Improving literacy and mathematics by ict-enhanced collaboration. *Computers & Education*, *99*, 68-80. https://doi.org/10.1016/j.compedu.2016.04.004.

Gravetter, F. J. Wallnau, L.B. (2016). *Statistics for the behavioral sciences.* Cengage Learning, Inc.

Gubbels, J., Netten, A., & Verhoeven, L. (2017). *Vijftien jaar leesprestaties in Nederland: PIRLS-2016.* Nijmegen: Expertisecentrum Nederlands, Radboud Universiteit.

Harmon, J. (2011). Research study finds iPad bolsters student reading and writing skills. *AASL Hotlinks*, *10*(7), 1-6.

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, *32(*2), 139 156. https://doi.org/10.1111/jcal.12123.

Heck, R.H., Thomas, S.L., & Tabata, L.N. (2014). *Multilevel and longitudinal modeling with IBM SPSS*. New York: Routledge.

Henderson, S., & Yeow, J. (2012). iPad in education: a case study of iPad adoption and use in a primary school. *Proceedings of the 45th Hawaii International Conference in System Science (hicss),* 78–87 IEEE. doi: 10.1109/HICSS.2012.390.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management, 23*(6), 723 – 744. doi: 10.1177/014920639702300602.

Hornstra, L., van der Veen, I., Peetsma, T., & Volman, M. (2015). Innovative learning and developments in motivation and achievement in upper primary school. *Educational Psychology*, *35*(5), 598-633. http://dx.doi.org/10.1080/01443410.2014.922164.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications.* New York: Routledge.

Huang, Y.-M., Liang, T.-H., Su, Y.-N., & Chen, N.-S. (2012). Empowering personalized learning with an interactive e-book learning system for elementary school students. *Educational Technology Research and Development*, *60*(4), 703–722. doi: 10.1007/s11423-012-9237-6.

Hutchison, A., Beschorner, B., & Schmidt-Crawford, D. (2012). Exploring the Use of the iPad for Literacy Learning. *The Reading Teacher*, *66*(1), 15–23. doi:10.1002/TRTR.01090.

Irwin, B. & Jones, N. (2011). *Pack your Laptop, We're off to School. The New Zealand Herald*. Retrieved from: http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=11191631

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The 2010 Horizon Report.* Austin, Texas: The New Media Consortium.

Jones, N. (2014). *School Gives an iPad to Every Pupil. The New Zealand Herald.* Retrieved from http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=11196246

Keeves, J. P. (1972). Educational environment and student achievement. *Stockholm Stud. Educ. Psychol*. 20: 1–309.

King, K., & Gurian, M. (2006). Teaching to the Minds of Boys. *Educational Leadership*, *64*(1), 56.

Koch, S. C., Müller, S. M., & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education*, *51*(4), 1795-1803. doi: 10.1016/j.compedu.2008.05.007.

Korat, O. (2010). Reading electronic books as a support for vocabulary, story comprehension and word reading in kindergarten and first grade. *Computers & Education*, *55*(1), 24–31. doi: 10.1016/j.compedu.2009.11.014.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of educational psychology*, *100*(1), 96. doi: 10.1037/0022-0663.100.1.96.

Krumsvik, R. J., Berrum, E., & Jones, L. Ø. (2018). Everyday Digital Schooling implementing tablets in Norwegian primary school. *Nordic Journal of Digital Literacy*, *13*(03), 152-176. doi: 10.18261/issn.1891-943x-2018-03-03.

LaRoche, S., Joncas, M., & Foy, P. (2016). Sample design in PIRLS 2016. *Methods and procedures in PIRLS*, 3-1.

Larson, L.C. (2010). Digital readers: The next chapter in e-book reading and response. *The Reading Teacher, 64*(1), 15–22. doi:10.1598/ RT.64.1.2.

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching*, *38*, 296–316.

Lim, C. S., Tang, K. N., & Kor, L. K. (2012). Drill and practice in learning (and beyond). *Encyclopedia of the Sciences of Learning*, 1040-1043. https://doi. org/10.1007/s40593-019-00186-y.

Lubienski, S. T., & Crane, C. C. (2010). Beyond free lunch: which family background measures matter? *Education Policy Analysis Archives*, *18*, 1-43.

Luke, D. (2004). *Multilevel modeling.* Thousand Oaks, CA: Sage.

Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *TRAMES: A Journal of the Humanities & Social Sciences*, *13*(1). doi: 10.3176/tr.2009.1.01.

Ma, X., Ma, L., & Bradley, K.D. (2008). Using multilevel modeling to investigate school effects. In A.A. O'Connell & D.B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59–110). Charlotte, NC: Information Age.

McConney, A., & Perry, L. B. (2010). Science and mathematics achievement in Australia: The role of school socioeconomic composition in educational equity and effectiveness. *International Journal of Science and Mathematics Education*, *8*(3), 429–452. https://doi.org/10.1007/s10763-010-9197-4

McFarlane, A., Triggs, P. & Wan, Y. (2008). *Researching Mobile Learning - Interim Report to BECTA Period: April - December 2007*, Bristol.

McPhee, I., Marks, L., & Marks, D. (2013). Examining the impact of the apple “iPad” on male and female classroom engagement in a primary school in Scotland. *ICICTE 2013 Proceedings*, 443-451.

Melhuish, K. & Falloon, G. (2010). Looking to the future: M-learning with the iPad. *Computers in New Zealand Schools: Learning, Leading, Technology, 22*(3), 1-16.

Mol, S., & Bus, A. (2011). Lezen loont een leven lang: De rol van vrijetijdslezen in de taal-en leesontwikkeling van kinderen en jongeren. *Levende Talen Tijdschrift*, *12*(3), 3-15.

Molenaar, I., van Campen, C. A. N., & Gorp, K. V. (2016). *Rapportage Kennisnet. Onderzoek naar Snappet; gebruik en effectiviteit.*

Nouri, J. (2016). The flipped classroom: for active, effective and increased learning especially for low achievers. *International Journal of Educational Technology in Higher Education*, *13*(1), 33. doi: 10.1186/s41239-016-0032-z.

NRC. (2019). *Laat de leraar niet buitenspel zetten door reken-app.* Retrieved from: https://www.nrc.nl/nieuws/2019/04/09/laat-de-leraar-niet-buitenspel-zetten-door reken-app-a3956276

OECD. (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills.* Paris: OECD.

OECD. (2005). *Are students ready for a technology-rich world? What PISA studies tell us.* Paris: OECD.

OECD. (2012). *Education at a Glance 2012.* OECD Publishing.

Osborne, J.W. (2017). *Regression and linear modeling: Best practices and modern methods.* Thousand Oaks, CA: Sage.

Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2017). Closing the gap: efficacy of a tablet intervention to support the development of early mathematical skills in UK primary school children. *Computers & Education*, *108*, 43-58. https://doi. org/10.1016/j.compedu.2017.01.011.

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2016). Comparing tablets and PCs in teaching mathematics: An attempt to improve mathematics competence in early childhood education. *Preschool and Primary Education*, *4*(2), 241-253. http://dx.doi.org/10.12681/ppej.8779.

PIRLS. (2016). *About PIRLS 2016.* Retrieved from: http://timssandpirls.bc.edu/pirls2016/international-results/pirls/about-pirls-2016/

PIRLS. (2018). *PIRLS 2016 Guide for the International Database.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). CA: Sage.

Saifi, S., & Mehmood, T. (2011). Effects of socio-economic status on students achievement. *International Journal of Social Sciences and Education*, *1*(2), 119-128.

Sanders, J. (2005). Gender and technology in education: A research review. In C. Skelton, B. Francis, & L. Smulyan (Eds.), *Handbook of gender and education*. London: Sage.

Sharples, M., Taylor, J., & Vavoula, G. (2007). A theory of learning for the mobile age. In R. Andrews & C. Haythornthwaite (Eds.), *The Sage handbook of Elearning research* (pp. 221–247). London: Sage.

Shuler, C., Winters, N., & West, M. (2012). *The future of mobile learning: implications for policy makers and planners.* Paris: UNESCO.

Schneeweis, N., & Zweimüller, M. (2014). Early tracking and the misfortune of being young. *The Scandinavian Journal of Economics*, *116*, 394–428. doi: 10.1111/sjoe.12046

Scholten, L., & Wolbers, M. H. (2019). Onderwijs van hoge kwaliteit. *Mens en maatschappij*, *93*(4), 375-406. https://doi.org/10.5117/MEM2018.4.004.SCHO.

Sheppard, D. (2011). Reading with iPads–the difference makes a difference. *Education Today*, *11*(3), 12–17.

Sieben, S., & Lechner, C. M. (2019). Measuring cultural capital through the number of books in the household. *Measurement Instruments for the Social Sciences*, *2*(1), 1. https://doi.org/10.1186/s42409-018-0006-0.

Smith, J. L., Morgan, C., & White, P. H. (2005). Investigating a measure of computer technology domain identification: A tool for understanding gender differences and stereotypes. *Educational and Psychological Measurement, 65*(2), 336–355. https://doiorg.proxy.library.uu.nl/10.1177/0013164404272486.

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science, 323*(5910), 122-124. doi: 10.1126/science.1165919.

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Strommen, E. F., & Lincoln, B. (1992). Constructivism, technology, and the future of classroom learning. *Education and urban society, 24*(4), 466-476.

Sun, L., Bradley, K. D., & Akers, K. (2012). A multilevel modelling approach to investigating factors impacting science achievement for secondary school students: PISA Hong Kong sample. *International Journal of Science Education*, *34*(14), 2107-2125. https://doi-org.proxy.library.uu.nl/10.1080/09500693.2012.708063.

Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis. *Computers & Education*, *94*, 252-275. https://doi.org/10.1016/j.compedu.2015.11.008.

Sutherland, P. A. (1992). *Cognitive development today: Piaget and his critics*. London: SAGE.

Tamim, R. M., Pickup, D., Borokhovski, E., Bernard, R. M., & El Saadi, L. (2015). Tablets for teaching and learning: A systematic review and meta-analysis.

Tangdhanakanond, K., Pitiyanuwat, S., & Archwamety, T. (2006). A Development of Portfolio for Learning Assessment of Students Taught by Full-Scale Constructionism Approach at Darunsikklialai School. *Research in the Schools*, *13*, 24-36.

Taylor, D., & Lorimer, M. (2003). Helping Boys Succeed. *Educational Leadership*, *60*(4), 68 70.

Timmermann, P. (2010). Is my iPad in my backpack? *Journal of Digital research & publishing, University of Sydney.*

Torppa, M., Eklund, K., Sulkunen, S., Niemi, P., & Ahonen, T. (2018). Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure reading and homework activity. *Journal of Research in Reading*, *41*(1), 122-139. https://doi-org.proxy.library.uu.nl/10.1111/1467 9817.12103.

Van de Bogart, W. (2012). *Child development issues related to Thailand's tablet computer policy within the ASEAN Community*.

Van de Gaer, E. V., Pustjens, H., Damme, J. V., & Munter, A. D. (2007). Impact of attitudes of peers on language achievement: Gender differences. *The journal of educational research, 101*(2), 78-90. https://doi-org.proxy.library.uu.nl/10.3200/JOER.101.2.7893.

Van der Meij, J., Kemps, J., Hoogland, I., & Rutten, N. (2015). Tablets in het basisonderwijs. *Een pilotonderzoek naar de verwachtingen van leerkrachten en de inzet van tablets in de lespraktijk.*

Vekiri, I., & Chronaki, A. (2008). Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & education, 51*(3), 1392-1404. doi:10.1016/j.compedu.2008.01.003.

Vosniadou, S. (2001). *How children learn. Educational practices series, 7.* Geneva, Switzerland: The International Academy of Education (IAE) and the International Bureau of Education.

Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives, 30*(3), 3-32. http://dx.doi.org/10.1257/jep.30.3.3.

Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian journal of educational research*, *47*(1), 21-41. doi: 10.1080/0031383032000033317.

Appendix

Appendix A: Results factor analysis

Table A1

Confirmatory factor analysis 1

Item	Loading
Short on instructional materials	*a*
Short on supplies	*a*
Sort on buildings	.421
Short on heating systems	*a*
Short on instructional space	.402
Short on technological staff	*a*
Short on audio-visual resources	.438
Short on computer technology	.465
Short on resources for students with disabilities	*a*
Short on teachers specialized in reading	.527
Short on software for reading	.720
Short on library resources	.660
Short on instructional materials for reading	.826
Eigenvalue	3.888
% variance explained	29.908

Note: Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.
a Factor loadings below .4 have been suppressed
b Factor extraction method is maximum likelihood
c Factor rotation method used is nonorthogonal rotation - Promax

Table A2

Confirmatory factor analysis 2

Item	Loading
Short on buildings	*a*
Short on instructional space	*a*
Short on audio-visual resources	*a*
Short on computer technology	.466
Short on teachers specialized in reading	.557
Short on software for reading	.749
Short on library resources	.648
Short on instructional materials for reading	.864
Eigenvalue	3.196
% variance explained	39.956

Note: Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.
a Factor loadings below .4 have been suppressed
b Factor extraction method is maximum likelihood
c Factor rotation method used is nonorthogonal rotation - Promax

Appendix B: Results reliability analysis

Table B1

Reliability statistics

	Correlation with total	Alpha if item deleted
General shortage – computer technology	.352	.797
Reading shortage – teachers specialized in reading	.467	.764
Reading shortage – computer software for reading	.723	.668
Reading shortage – library resources	.511	.749
Reading shortage – instructional materials	.740	.677
Cronbach's alpha scale	.777	
N	5	

Note: Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.

Appendix C: Critical values chi-square distribution

	Proportion in Critical Region				
df	0.10	0.05	0.025	0.01	0.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.53	20.09	21.96
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
11	17.28	19.68	21.92	24.72	26.76
12	18.55	21.03	23.34	26.22	28.30
13	19.81	22.36	24.74	27.69	29.82
14	21.06	23.68	26.12	29.14	31.32
15	22.31	25.00	27.49	30.58	32.80
16	23.54	26.30	28.85	32.00	34.27
17	24.77	27.59	30.19	33.41	35.72
18	25.99	28.87	31.53	34.81	37.16
19	27.20	30.14	32.85	36.19	38.58
20	28.41	31.41	34.17	37.57	40.00

***Figure 1*. Original caption of the figure. Adapted from "Statistics for the behavioral sciences," by F.J. Gravetter and L.B. Wallnau, 2016, Cengage Learning, Inc.**

Appendix D: Calculations of model fit indices

Table D1

Model comparison in terms of fit indices

	Null model	Model 2	Model 3	Full model
ICC [a]	.127			
-2LL [b]	45473.379	43083.512	31483.940	31843.895
AIC [c]	45479.379	43099.512	31863.940	31865.895
R_1^2 [d]		.106	.138	.138
R_2^2 [e]		.333	.461	.461

Note: Data generated from the Progress in International Reading Literacy study (PIRLS) of 2016.
[a] ICC = intraclass coefficient
[b] -2LL = Log Likelihood
[c] AIC = Akaike information criterion
[d] R_1^2 = Explained variance at the student level
[e] R_2^2 = Explained variance at the school level

The null model

ICC

The ICC for this model = 402.671 / (402.671 + 2758.479) = .127.

Model 2

Log likelihood

For the null model the -2LL is χ^2 = 45473.379 and for the second model χ^2 = 43083.512. Hence, χ^2_{change} = 45473.379 – 43083.512 = 2389.867. DF_{change} is 8 – 3 = 5. The critical value for a DF_{change} of 5 is 11.07. Since 2389.867 is higher than 11.07, it can be concluded that the second model fits significantly better than the null model.

AIC

The AIC of model two is lower than that of the null model (43099.512 vs. 45473.379), indicating that the second model fits the data better.

Explained variance

- The explained variance at the student level: (2758.479 - 2465.970) / 2758.479 = .106
- The explained variance at the school level: (402.671 - 267.909) / 402.671 = .333

Model 3

Log likelihood

χ^2 in the second model = 43083.512 and for the third model χ^2 = 31483.940. This results in χ^2_{change} = 43083.512 – 31483.940 = 11599.572, with a DF_{change} of 10 – 8 = 2. The critical value for a DF_{change} of 2 is 5.99. Since χ^2_{change} is higher than this value, the third model is a better fit than the second model.

AIC

The AIC reports 31863.940 which is lower than the AIC of the second model, hereby suggesting that the third model is the better model.

Explained variance

- The explained variance at the student level: (2758.479 - 2378.379)/ 2758.479 = .138
- The explained variance at the school level: (402.671 - 217.089)/ 402.671 = 0.461

Model 4

Loglikelihood

For this full model, χ^2 = 31843.895 and for the third model χ^2 = 31483.940. Thus, χ^2_{change} = 31483.940 – 31843.895 = -359.955, with a DF_{change} of 11 – 10 = 1. The critical value for a DF_{change} of 1 is 3.84. Since this is lower than the critical value, it can be concluded that the third model is a better model than the third model.

AIC

The AIC indicates that the third model is the better model, as there is only a minor difference between the AIC of the third model and that of the full model (model 3 = 31863.940; full model = 31865.895).

Explained variance

- The explained variance at the student level: (2758.479 - 2378.325)/ 2758.479 = 0.138
- The explained variance at the school level: (402.671 - 217.158)/ 402.671 = 0.461