IMPLEMENTING TABLETS IN NORWEGIAN PRIMARY SCHOOLS: EXAMINING OUTCOME MEASURES IN THE SECOND COHORT

Rune J. Krumsvik, University of Bergen
Erling Berrum, Rambøll Management
Lise Øen Jones, University of Bergen
Ingrid P. Gulbrandsen, Rambøll Management
Norway

Abstract

This study examines the implementation of tablets in primary schools in Norway. The outcome measures in the study are external for the intervention and are recorded data from national tests (National reading, arithmetic and English Tests, Classes 5, 8 and 9; National Mapping Tests for reading and arithmetic, Classes 1–3; and the 2014–2017 National Pupil Survey). The entire study (N=15, 708) relies on an explanatory, sequential mixed-methods design (Fetters, Curry, & Creswell 2013), and in this study we examine the quantitative effects of this implementation. The results indicate that the impact of tablets on pupils' school achievement varies. It seems that tablets contribute more positively to boys' school achievements than to girls' school achievements. However, we cannot rule out that a grade effect may also have an impact on the results, and we therefore request that the results be read with this reservation.

Keywords: tablets, digital schooling, implementation, primary school, Norway

1. Introduction

This article examines the second cohort of the trailing research in the Municipality of Bærum's (2015) *Everyday Digital Schooling* tablet project, which examines outcome measures regularly through our longitudinal research design. The first study examined the first nine months of this project (Krumsvik, Berrum, and Jones, 2018). This second study examines the next 24 months of the project period. These two first studies are the first large-scale effect studies of the implementation of tablets in Norwegian primary schools where the outcome measures are external for the intervention, as recommended by, for example, Cheung and Slavin (2013). This means that the learning outcome in this study is the combined result of national tests, the National Mapping Tests and the National Pupil Survey (administered by The Norwegian Directorate of Education and Training).

The aim of introducing tablets as a primary learning aid for all pupils at all stages at the pilot schools was to improve the academic and personal outcomes acquired by the pupils from their schooling. Investing in tablets had two objectives: to challenge teachers to develop and change their own teaching

and working practices wherever possible, and to help with the provision of better learning for pupils. However, to avoid Cheung and Slavin's (2013) critique concerning educational technology studies using measures designed by the researchers themselves, we applied external outcome measures (registry data). In this part of the trailing research, the outcome measures in the study are external for the intervention and are recorded data from National Tests (National reading, arithmetic and English tests, Classes 5, 8 and 9, National Mapping Tests for reading and arithmetic, Classes 1–3, and the 2014–2016 National Pupil Survey) in a municipality in Norway. In this second cohort of the trailing research, we only examine the quantitative effects of this part of the implementation. The paper first presents a conceptual framework and the methodology of the study, followed by the results and a discussion of the study's main findings.

2. Conceptual framework

2.1. Literature Review

Norwegian schools are implementing tablets in schools to an increasing degree, and there seems to be a need for more research within this area to examine how this implementation affects pupils' learning processes (OECD 2008; Krumsvik, Egelandsdal, Sarastuen, Jones, & Eikeland, 2013). There are a limited number of large-scale research studies within the application of this kind of tablet technology for educational purposes. More research is therefore needed within this area, especially since we know that throughout the world there are initiatives at various policy levels regarding the implementation of tablets in schools.

Norway has had a high technology density both in homes and in schools during the last 10 years, and it is therefore interesting to examine how tablets affect school achievement variables. This is also related to the present national curriculum (Kunnskapsdepartementet, 2006) and the upcoming national curriculum (The Ministry of Education 2017; The Norwegian Directorate of Education and Training, 2018a), which both highlight digital skills and digital competence among pupils in school.

A recent doctoral thesis from Norway by Kongsgården (Kongsgården, 2019; Kongsgården & Krumsvik, 2016; Kongsgården & Krumsvik, 2019), shows that the implementation of tablets in schools is a complex process with both new educational possibilities and pitfalls. The study shows that tablets play a certain role in the learning process, especially in the achievement of learning goals and access to the Internet. However, there are clear differences in how pupils use tablets in their learning processes. In particular, there is a difference between primary and secondary school. Kongsgården's study (Kongsgården, 2019) also indicates that a teaching design that includes educational technology contributes to an increase in learning outcomes. Through the teacher's didactical choice, there is evidence that the teacher, by creating a learning community focusing on assessment for learning and technology, establishes flexible and transparent learning processes that develop the pupils' self-regulation. The study shows that the critical success factor is the teacher

and his or her ability to create a teaching plan where the use of technology is justified by didactic choices and not vice versa (Kongsgården, 2019; Kongsgården & Krumsvik, 2016; Kongsgården & Krumsvik, 2019).

Another PhD study from Norway examines the effect of adaptive learning technologies (ALT) and the use of tablets (Moltudal, Høydal, & Krumsvik, 2019) in grades five to seven (10-12 years of age) in mathematics. The findings of the study indicate that the use of ALT at the upper primary level contributed positively to basic pupil learning in mathematics (ES = 0.39, P = 0.001). However, the study also indicates an intertwined relationship among learning, motivation, and volume training, especially for pupils learning new mathematical concepts. However, successful implementation requires that teachers have expertise in classroom management. It also shows that one of the main educational challenges lies in changing teachers' traditional practice by implementing a digital didactic method that provides the teacher with a greater understanding of digital homework as a measure *for*, and opportunity to better understand where pupils are *during*, the learning process.
Tamim, Borokhovski, Pickup, Bernard & El Saadi (2015a) carried out a systematic review of current government-supported tablet initiatives around the world, in order to understand more of the educational basis and underlying principles in general. This review concluded "that the majority of these initiatives have been driven by the tablet hype rather than by educational frameworks or research-based evidence" (p. 9).

To a certain degree, Escueta, Quan, Joshua, and Oreopoulos (2017) find some of the same tendencies in their evidence-based review of educational technology in general. They find that it is not enough to provide students with access to technology – it has to be based on a reflective pedagogical teaching design.

Fairlie and Robinson (2013) revealed much of the same when they examined the effects of home computers on academic achievement among schoolchildren. They concluded that "we find no evidence that home computers had an effect (either positive or negative) on any educational outcome, including grades, standardized test scores, or a host of other outcomes" (p. 234). From these three studies (and also from earlier meta-analysis as e.g., Tamim, Bernard, Borokhovski, Abrami & Schmid (2011), we can see that *access* to technology is not enough – it seems to be a consensus in the research community that technology has to be closely attached to well-founded pedagogy and didactics. So, what do we know from recently published meta-analyses about tablets and mobile technology in pedagogical settings?

A meta-analysis by Sung, Chang, & Liu (2016) finds that "the overall mean effect size for learning achievement…was 0.523, meaning that learning with mobiles is significantly more effective than traditional teaching methods that only use pen-and-paper [*sic*] or desktop computers" (p. 257). For tablet PCs, they find a specific effect size of 0.615. Sung et al. (2016) also state that if we compare these effect sizes with Kulik and Kulik's (1991) and Tamim et al.'s (2011) meta-analyses of the difference between using computers and not using

computers in education (effect size between 0.30–0.35), some of the reason for these improved effects might be attached to the affordances that specific tablet and mobile technology give. However, Sung et al. (2016) emphasise that more research is needed to examine such issues.

Tamim, Borokhovski, Pickup, Bernard & El Saadi (2015b) carried out a meta-analysis of 68 studies based on 27 quantitative studies and 41 qualitative research studies, and concluded that "findings from the current meta-analysis indicate a moderate strength average effect size for the impact of tablets and smart mobile devices on student outcome measures" (p. 38).

These two meta-analyses are up to date, give some promising results, and indicate that tablets represent a type of hardware with affordances other than those of traditional computers. However, these are preliminary tendencies, and we need more research into the affordances tablets might or might not give. Concerning literacy more specifically, Genlott and Grönlund (2016) examined the effects of the "Write to Learn" (WTL) method. The results showed that the WTL group achieved the best results, and they concluded that access to technology is not enough; information communication technologies (ICT) have to be included in both didactical and pedagogical elements in instruction.

In their meta-analysis of the effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms, Cheung and Slavin (2013) find only a positive, modest effect of d=0.15. In another meta-analysis examining how features of educational technology applications affect student reading outcomes, they also find positive, modest effects of d=0.16 (Cheung & Slavin, 2012). They explain that high quality studies (included in their meta-analysis) within educational technology give a lower effect size than do studies with methodological weaknesses (excluded from their meta-analysis).

On the basis of this literature review, we find that despite the existence of some international research concerning tablets (and other types of educational hardware) in schools, we have very little research knowledge about how the large-scale implementation of tablets affects pupils' learning outcomes in Norway. Our trailing research is therefore positioned towards this gap, and will provide empirical data as related to our research questions.

2.2. Theoretical Framework

Certain theoretical discussion is related to whether it is the educational technology (e.g. tablets) by itself that affect learning or whether it is the teaching method, teacher and other factors. Such debates have been going on since the 1980's and are still debated in today's research communities. However, Cheung and Slavin (2013) provide a certain "middle way out" solution:

> Though it may be theoretically interesting to ask whether the impact of technology itself can be separated from the impact of particular applications, in practice, technology, content, and method are often intertwined and cannot be separated. As is the case for many

> educational interventions with many components, currently available technology applications can be seen as packages of diverse elements and evaluated as such. If a particular combination of hardware, software, print materials, professional development for teachers, and other elements can be reliably replicated in many classrooms, then it is worth evaluating as a potential means of enhancing student outcomes. Components of effective multi-element treatments can be varied to find out which elements contribute to effectiveness and to advance theory, but it is also of value for practice and policy to know the overall impact for students even if the theoretical mechanisms are not yet fully understood. (p. 92)

Thus, this paper has no ambitions to develop new theory, but to apply theory as Leedy and Ormrod (2005, p. 4) describe it: "A theory is an organized body of concepts and principles intended to explain a particular phenomenon". The theoretical framework for the entire study underpins the research questions (and are not an analytical framework). The theoretical framework refers to the theories of Piaget (1967) and Vygotsky (1978), where tablets are related to both knowledge construction and collaborative learning, and linked to student-centred and group-based teaching design. Educational technology (like tablets), as it appears today in Bærum schools with its distinctive feature of digital tools, relates especially to more recent socio-cultural perspectives on learning (Wertsch, 1998; Cole, 1996; Säljö, 2005, 2017; Stahl, 1993; Lave & Wenger, 1991; Wenger, 1998) as a mediating artefact. The socio-cultural perspective emphasises the point that learning is constructed in interaction with other people and mediating artefacts, which has a significant focus on the basic thinking in the "Digital everyday school" school development project. James Wertsch states that such new kinds of mediation and mediated artefacts can give new possibilities and the experience of "…how the introduction of novel cultural tools transforms the action" (Wertsch, 1998, p. 42). The use of tablets for learning purposes also relates to Richard Mayer's (2010) *Multimedia Learning Theory* where he describes learning *with* technology, such as situations wherein technology is used for the purpose of promoting learning, and is concerned with the human *construction of knowledge* as a framework for learning.

However, tablets are a type of hardware that can be applied in numerous ways, and it is important to understand the affordances of such technology and the context of use. This is based on the fact that there are several similarities between ICT for entertainment use and educational technology for use in school, and sometimes it is hard to distinguish the two. However, educational technology is developed especially for educational purposes, while ICT consists of a myriad of technologies such as social media, mobile phones, wireless broadband, PCs, and so on, which are developed first and foremost for everyday life (and not specifically for educational purposes). Tablets can be used in both contexts, but in this study we examine tablets as an educational technology with certain affordances for teaching and learning in school contexts. Cheung and Slavin (2013) state that educational technology has a variety of definitions in the literature; in this paper educational technology refers to the use of tablets in school settings for educational

purposes to support learning process and learning goals. Thus, our theoretical underpinning for the study is also attached to *digital didactics*. This concept was introduced by Krumsvik (2008) and was further examined in subsequent studies (Krumsvik, 2009a; Krumsvik, 2009b; Almås & Krumsvik, 2008; Krumsvik, 2012). Similar to the later digital didactic models of, for example, Jahnke, Bergström, Mårell-Olsson, Häll, and Kumar (2017), this digital didactic model focuses on the most relevant elements teachers need to consider in the digitalised school with the awareness that "… adding 21st-century technologies to 20th-century teaching practices will just dilute the effectiveness of teaching" (Organisation for Economic Co-operation and Development [OECD, 2015, p. 5).

Another element to consider (which has both theoretical and methodological implications) is that:

> Many evaluations of technology applications suffer from serious methodological problems. Common problems include a lack of a control group, limited evidence of initial equivalence between the treatment and control group, large pretest differences, or *questionable outcome measures*. In addition, many of these reviews included studies that had a very short duration. Unfortunately, studies with poor methodologies tend to report much higher effect sizes than those with more rigorous methods (…), so failing to screen out such studies inflates the average effect sizes of meta-analyses. (Cheung & Slavin, 2013, p. 92, our italics)

On this basis, the outcome measures in this study lies outside the intervention (registry data). The coherence among pupils' knowledge construction and collaborative learning linked to student-centred teaching design in schools (attached to sociocultural theory), learning with technology (tablets) attached to multimedia learning theory, and teachers' pedagogical practices (in relation to digital didactic) underpins the research questions of the study, which in the first cohort were:

1. To what extent does the implementation of tablets affect learning outcomes in schools in Bærum Municipality (where the outcome measures are recorded data such as National Mapping Tests, National Tests and the National Pupil Survey)?
2. To what extent does the implementation of tablets affect social enjoyment and learning environments in schools in Bærum Municipality (based on the National Pupil Survey)?

To be able to examine these same variance research questions in the second cohort, we have chosen trailing research and mixed method research, described below.

3. Methodology

The research design made use of trailing research (Finne, Levin, & Nilssen, 1995) and mixed method research (Fetters et al., 2013), which involved combining different methods and data sources. To be able to answer the research questions in this study, we have chosen to design this study as an

explanatory, sequential mixed-methods design (Fetters et al., 2013). We follow the *staged approach*, which means that data are reported in stages and published separately. In this article (the second cohort), we therefore only report the quantitative effect analysis which is based on existing recorded data. The effects of the learning results are measured by using the following data sources:

1. National reading, arithmetic and English tests, classes 5, 8 and 9 from 2014-2017.
2. National Mapping Tests for reading and arithmetic, classes 1–3 from 2015-2017.
3. The 2014–2017 National Pupil Survey.

We have obtained the results of the National Tests from the Norwegian Directorate for Education and Training's school portal, and the results of the National Mapping Tests have been provided by the Municipality of Bærum. Our two endpoints in this respect are based on class levels, divided according to gender and test type. Data from the national arithmetic and English tests have been taken from 2014 to 2016, since there is no comparable data available prior to 2014. The reading test is nevertheless included in our analysis, but with the reservation that changes have been made to the scale, so that the comparison cannot be made beyond 2016. However, this should not be a problem since the comparison is only made up to 2016. As regards the Mapping Tests, two respective tests were conducted in reading and arithmetic between 2014 and 2016.

Our third and final endpoint is social enjoyment and learning environments. This has been gathered from the National Pupil Survey. The National Pupil Survey focuses on how pupils perceive their learning environment at school, how motivated they are, their social well-being at school, if they experienced any bullying, how they experience the teachers, and so on. The results of the National Pupil Survey have also been obtained from the Norwegian Directorate of Education and Training's[1] school portal, based on class levels and divided by gender. Our basis includes the various indicators defined by the Directorate as being relevant for pupils' learning environments. We used data from the National Pupil Survey covering 2013 to 2015. No data for 2016 was available in the school portal when our analysis was carried out.

4. Quantitative Results

This section presents the quantitative surveys that have been made and the findings that emerge from these. We will present the analyses of our effect analyses, which are based on the last available registry data. Here we investigate the effect of the introduction of tablets on pupils' learning outcomes (in basic skills) and learning environments. The three effect measures analysed are the results of the National Tests in the fifth and ninth grades, the National Mapping Tests first to third grade, and the results from

[1] More information here: https://skoleporten.udir.no/

the National student survey (The Norwegian Directorate of Education and Training, 2018) in the seventh and tenth grades.

4.1. Effect Analyses

The purpose of the effect analyses is to investigate the effect of introducing tablets into pupils' learning exchange and learning environment. Then, pupils' learning outcomes and learning environment are compared with schools where tablets have not yet been introduced for all pupils.

The impact on learning outcomes is measured using the following data sources:

1. National tests in reading, mathematics, and English in the fifth, eighth and ninth grades.
2. National Mapping Test in reading and mathematics in first through third grade.

The results from the National tests are taken from the website of the Norwegian Directorate of Education and Training's (2018), "Skoleporten", as well as from the results of the national survey tests which we received from Bærum Municipality. Our two effect measures here are based on grade level, divided by gender and type of test. For the mapping tests, two tests are carried out in reading and mathematics, respectively.

The impact on pupils' learning environment is measured using collected data from the National Student Survey (The Norwegian Directorate of Education and Training, 2018) in seventh and tenth grades, based on grade and divided by gender. Furthermore, we use the different indicators that the Directorate of Education has defined as relevant to pupils' learning environment.

All three effect targets are linked with data at the school level from the "Primary School Information System" (GSI) in addition to socioeconomic indicators for the 24 children's schools in Bærum municipality.

4.1.1. Description of the Sample as the Basis for Effect Analyses

Table 1 below describes the pupils in Pilot 1 and Pilot 2 schools, as well as the pupils at other schools, where we investigate whether or not there are differences between schools that have used tablets and schools that have not.

Table 1
Description of the Pilot Schools and Non-pilot Schools

	Pilot 1	Pilot 2	Non-pilot schools
Number of schools (total)	5	10	29[a]
Number of pupils (total)	1,743	4,395	9,570
Percentage of secondary schools	40 %	30 %	31 %
Percentage of schools above 400 pupils	20 %	60 %	34 %
Average number of pupils per year[b]	15.3	16.4	13.4
Average number of assistant hours per pupil	10	8	23
Sociodemographic variables:[c]			
Percentage with low income (b. 50% median)	7.7 %	7.1 %	7.4 %
Percentage with low or no education	18.5 %	16.3 %	17.2 %
Percentage of social help recipients	2.0 %	1.1 %	1.6 %
Percentage with immigrant background	18.0 %	13.5 %	16.2 %

Note: There are no significant differences between group schools and other schools. The significance is tested by a two-tailed independent T-test with equal variance of 10 %, 5 %, and 1 % significance level.

[a]The 10 group schools from group 2 were taken out of the control group when they introduced tablets in August 2016 and therefore cannot act as a control group for an after-survey survey in 2017. [b]There is a significant difference between group 2 schools and other schools in the variable average number of students per year at a 10 % significance level. There are otherwise no significant differences between group schools and other schools on the other variables. [c]Source: Indicators from 2011 in nine areas in Bærum calculated by Statistics Norway. The distribution between the schools is made by the Municipality of Bærum. For some schools, a percentage distribution has been developed between several areas.

Findings

Pilot 1 schools do not differ significantly from other schools in Bærum. In the socioeconomic parameters, there are also no statistically significant differences between Group 1 schools and other schools. As described in the previous report, one should be careful when drawing conclusions based on the socioeconomic variables, as they are from 2011. At the same time, the pupil base in the surrounding area is expected to be relatively constant as the school district changes only marginally each year. In the analysis, the indicators are used only to test the robustness of the results in comparative analyses, and not as an independent analysis.

Group 2 schools differ from other schools by having a slightly lower proportion of secondary schools, larger schools, more students per year, and fewer assistant hours per student. However, these differences are on the whole not significant.

In the socioeconomic parameters, we see that Group 2 schools are in an area with a lower proportion of children with immigrant background than are the other schools (the opposite of what we see for Group 1 schools). However, there are no statistically significant differences between the school groups in any of the socioeconomic parameters.

The parameter showing the greatest variation between the three school groups is "Number of students per year". Here, the other schools have the lowest average. This could potentially contribute to better student outcomes for these students. However, we have taken this into account through our difference-in-difference analytical approach (see 4.1.2).

4.1.2. On Method and Identification of Effect

The effect analysis is performed with a difference-in-difference approach in a simple average analysis and a more advanced fixed-effect regression analysis. In a simple "diff-in-diff" analysis, the average difference between the five Pilot schools and all other schools in Bærum is considered before the introduction of tablets. This is compared with the difference between group schools and all the other schools in Bærum after the introduction of tablets. Figure 1 below illustrates the difference-in-difference approach in our study.

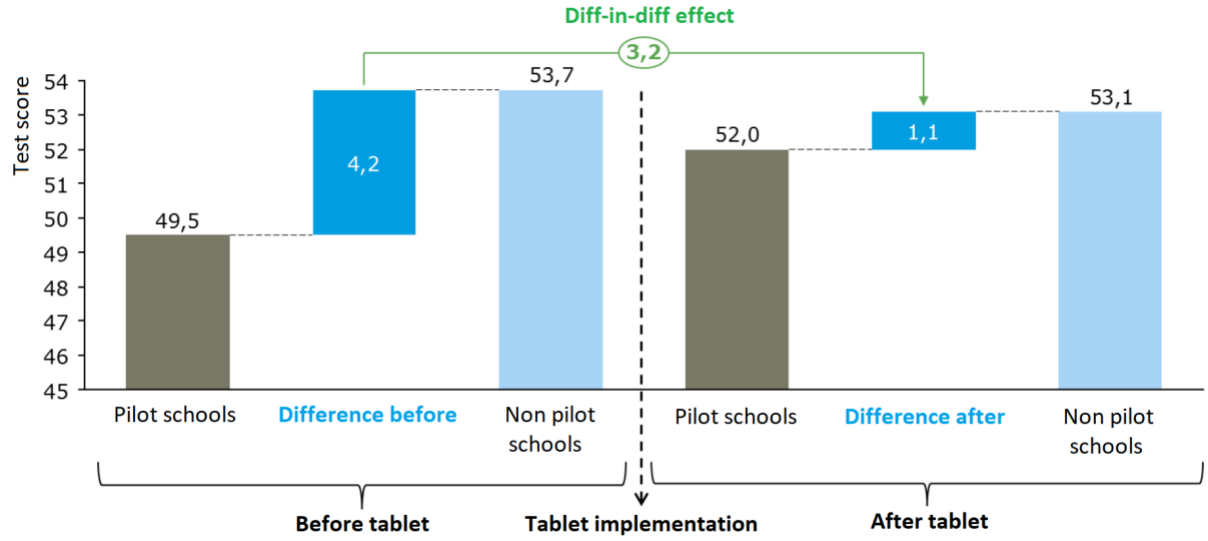

Figure 1. Illustration of the difference-in-difference approach. The green bubble is the estimated effect of the introduction of tablets.

Using a diff-in-diff approach in a more advanced fixed-effect regression analysis, as you can check for time constant variables at the school level. This means variables that do not change over the years - such as school size, geographical location, and organisation - will be checked for. In addition, the method takes into account unobservable characteristics that are constant over the years, such as school culture, student basis (assuming student base is not changing), and the like.

4.1.3. Reservations and Uncertainty in the Analysis

In diff-in-diff analyses (both simple and fixed-effect analysis), it is assumed that schools would have developed equally if the pilot schools had not introduced tablets. This assumption is necessary, as in a diff-in diff analysis the pilot of schools without intervention defines the counterfactual situation of schools that have introduced tablets. That is, after taking into account the different starting points of the school before the introduction of tablets, they are expected to have the same development over the years in the national tests, national mapping tests, and the National Student Survey. This is a strict

assumption, and it cannot be tested in the data we have available. Therefore, in the interpretation of the results, it should be noted that there may be cases where Group 1 schools without the introduction of tablets could still have developed as they did. One way to approach this strict assumption is to include variables that describe pupils' individual backgrounds. As we have not had access to such data, we have also not had the opportunity to take this information into account in the analysis.

In addition to the strict assumption of development, another uncertainty occurs in the form of a "grade effect". By grade effect, it is believed that the analysis is based on the comparison of students in a single grade, for example, in fifth grade, with the subsequent graduation of students in fifth grade. In other words, the same students are not followed. This implies that there may potentially be students who overall are better or worse, contributing to a proven effect of tablets, and not the characteristics of the tablets themselves. The grade effect can be tested by following a student group over two grades (for example from first to second grade), thus evaluating whether the tablet changes the results in the same student group.

This also means that the results cannot be generalised to other schools or municipalities. Furthermore, we have an analysis of measurable effects, which means that the analysis does not capture potential effects on learning beyond the measurable indicators. All results must therefore be seen in the light of these reservations.

4.1.4. Identification of Effects

The chart below (Figure 2) shows an overview of when the group schools introduced tablets. The overview also shows when the various impact targets were collected at a national level. Furthermore, the grey areas mark the years used as before and after measurements.

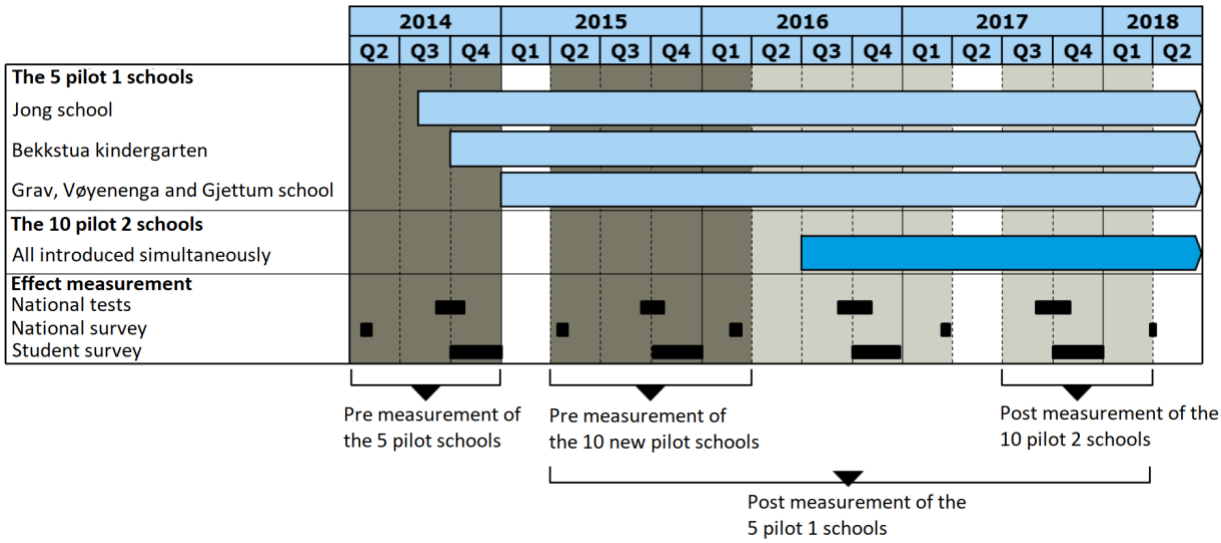

Figure 2. Overview of the introduction of tablets and the three effect measurements.

The effect measurements from 2014 and 2015 are used as preliminary measurements for Group 1 and Group 2 schools, respectively. However, it must be noted that the pre-measurement of the National Student Survey and the National Tests for Jong school and Bekkestua primary school may be influenced by the fact that the schools in question introduced tablets already in autumn 2014. However, state surveys in 2014 and 2015 qualify as preliminary measurements for all schools, as they were collected in the spring of the same year.

The reason 2013 data is not used in the National Tests for Group 1 schools is that the National Tests in 2013 are not comparable with data from 2014 and later. For the student survey, however, 2013 can be used as a measure for Group 1 schools. Nevertheless, the measurements from 2014 are used to see the three analyses in one. As a reassessment, data are used from 2015, 2016, and 2017.

4.1.5. Results from National Tests in Primary School

Results will be divided so that the results of the national samples are described first. Then the results of the surveying tests are presented, and finally the results from the student survey. In conclusion, a brief summary of the results follows.

Effects for Group 1 in Fifth Grade (Analysis 1)

Table 2 shows the average test results for national tests in reading, arithmetic and English for all children, boys and girls. A positive number in the Effect column on the right indicates that Group 1 schools have developed favourably compared to other schools in Bærum after the introduction of tablets. The analysis was completed in 2017, i.e., it reports on the effect for 2017. In addition, the results of the previous report are included in the first column in order to compare short-term and longer-term effects.

Table 2
Difference-in-difference Analysis of Fifth Grade Test Results (Pilot 1 Schools)

		Effect (2015,2016)	Before tablet (2014)		After tablet (2017)		Effect (2017)
		Diff-in-diff	Pilot schools	Non-pilot schools	Pilot schools	Non-pilot schools	Diff-in-diff
All	Reading	3,4	-	-	-	-	-
	Arithmetics	3,2	49,5	53,3	51,5	51,3	4,0
	English	3,7	50,5	53,0	55,0	51,5	6,0**
Boys	Reading	4,3	-	-	-	-	-
	Arithmetics	5,9*	49,5	54,8	52,5	52,6	5,1
	English	4,7*	49,5	53,7	56,0	52,3	7,9***
Girls	Reading	1,7	-	-	-	-	-
	Arithmetics	0,3	49,5	52,1	49,5	50,8	1,4
	English	3,1	51,0	53,1	53,5	51,3	4,3
Number of schools [a)]			**2 Schools**	**16 Schools**	**2 Schools**	**16 Schools**	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10 %, 5 %, and 1 % significance level. If a number does not include asterisks, there is no statistical difference.

[a] Bekkestua Primary School is not included in the analysis, as at the time of measurement it did not have its own fifth grade.

*with 90% certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

National tests in reading cannot be compared after 2016, as changes have been made to the scale of this test. Therefore, the result for reading is omitted from the analysis, as all the measurement for reading takes place after 2017.

In general, the impact of tablets has increased since the measurements collected in 2015 and 2016.

The effect of introducing tablets is significantly positive for *boys* in fifth grade in English (as in 2015/2016). Furthermore, the effect is also positive and significant for *all children* in fifth grade in English, when the effect is measured in 2017. For *girls*, we cannot say with statistical certainty that a change has occurred. If a change is to be found in the latter group, the results indicate that the change is likely to be positive.

The fifth-grade boys also had a significant positive effect in the use of tablets in mathematics measured in 2015/2016. This effect is no longer significant in 2017.

We also conducted a similar analysis for the three levels of mastery in arithmetic, reading, and English (data is available upon request). In general, the proportion of students in third grade in English rises significantly more for pilot schools than other schools after the introduction of tablets. It also results in a significant negative effect in Level 2 (albeit trend of positive performance), as a large proportion of Level 2 students pass to Level 3.

Effects for Group 2 in Fifth Grade (Analysis 1)

Table 3 shows the average test results for national tests in mathematics, reading, and English for all children, boys and girls. A positive figure in the Effect column on the right indicates that Pilot 2 schools have developed more positively than the other schools in Bærum after the introduction of tablets. Both 2016 and 2017 are included in the aftermath, which means that the measured effect is an average of the effects in 2016 and 2017.

Table 3
Difference-in-difference Analysis of Fifth Grade Test Results (Pilot 2 Schools)

		Before tablet (2014,2015)		After tablet (2016,2017)		Effect
	Basic skill	**Pilot schools**	**Non-pilot schools**	**Pilot schools**	**Non-pilot schools**	**Diff-in-diff**
All pupils	Reading [a)]	53,5	53,2	53,4	52,1	1,0
	Arithmetic	54,4	53,1	53,3	51,9	0,1
	English	53,0	53,0	52,9	51,9	1,0
Boys	Reading [a)]	52,3	53,3	53,3	51,7	2,5
	Arithmetic	55,4	54,3	54,2	52,9	0,2
	English	53,8	53,5	53,5	53,0	0,2
Girls	Reading [a)]	53,7	53,7	53,4	52,3	1,1
	Arithmetic	53,3	51,9	52,1	51,3	-0,7
	English	52,4	52,8	52,0	51,2	1,3
Number of schools		**7 schools**	**16 schools**	**7 schools**	**16 schools**	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10 %, 5 %, and 1 % significance level. If a number does not include asterisks, there is no statistical difference.

[a]National tests in reading cannot be compared beyond 2016, as changes have been made to the scale of this test. The sample is therefore not included in this type of sample in 2017.

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

It is considered that national tests in reading cannot be compared to 2016, since the reading for the reading exam consists only of 2016. This is also described in the note below the table.

There are no statistically significant effects to be found for pilot 2 schools as compared to other schools measured in terms of the national fifth-grade tests. This corresponds to the fact that we did not find any effect for pilot 1 schools at this time (i.e. after a relatively short period of time).

Effects for Group 2 in Fifth Grade (Analysis 2)

The fixed effect analysis in Table 4 (group 2) reinforces the results in the difference-in-difference analysis from Table 3 (group 2), where we do not find positive significant effects for all students or any of the two gender groups. At the same time, note that the effect in reading for boys in the fifth grade is significantly positive, albeit as a short-term effect, as the effect of introducing tablets on reading skills is only measured in 2016 (see point below). This means that in 2017 we cannot say with statistical certainty that there has been a positive change in the development of students' reading skills.

Table 4
Difference-in-difference in Fixed Effect Regression Analysis in Fifth Grade (Pilot 1 schools)

	All pupils		Boys		Girls	
	Arithmetic	English	Arithmetic	English	Arithmetic	English
Group school	0	6,3	-2,7	1	4	-8,7
After implementation	-2,1***	-3,2***	-2,2**	-1,5*	-2,1**	-2,3*
Effect of tablet (2017) [a)]	4,3**	3,2*	5,7	9,1***	3	2
Large schools [b)]	-0,7	8,1***	-1,2	-2,3	-2,7	5,1
Number of pupils per year	-0,4	-1,5***	-0,3	-0,3	-1,1**	-0,7
FE, School level	Ja	Ja	Ja	Ja	Ja	Ja
Explanatory power (R^2)	0,93	0,88	0,84	0,81	0,89	0,66
Number of obs.	36		36		37	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10 %, 5 %, and 1 % significance level. If a number does not include asterisks, there is no statistical difference.

[a]The effect of tablets is an interaction between a dummy variable to be the intervention school and dummy variable to be after the implementation. I.e. the effect is calculated by a difference-in-difference approach. [b]A big school is defined as a school with 400 students or more.

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

The national test in reading cannot be compared to 2016, since the reading for the reading exam consists only of 2016. It is also described in the note below the table. For the other national tests (Arithmetic and English), both 2016 and 2017 have been included in the survey.

The fixed effect analysis has taken into account time-constant characteristics at school level, as well as school size and number of students per year.

4.1.6. Results from National Tests at Secondary School

Table 5 shows the average test results for national tests in arithmetic and reading for all children, boys and girls, in ninth grade for Pilot 1 schools and other schools.

Table 5
Difference-in-difference Analysis of Average Test Results, Ninth Grade (Group 1)

		Effect (2015,2016)	Before tablet (2014)		After tablet (2017)		Effect (2017)
		Diff-in-diff	Pilot schools	Non pilot schools	Pilot schools	Non pilot schools	Diff-in-diff
All	Reading[a)]	-1,7	-	-	-	-	-
	Arithmetic	-2,1	57,5	57,1	56,0	57,6	-1,9
Boys	Reading[a)]	0,2	-	-	-	-	-
	Arithmetic	-1,9	58,0	57,9	58,0	58,9	-1,0
Girls	Reading[a)]	-3,0	-	-	-	-	-
	Arithmetic	-2,7	56,5	56,3	54,5	55,9	-1,6
Number of schools			**2 schools**	**7 schools**	**2 schools**	**7 schools**	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10 %, 5 %, and 1 % significance level. If a number does not include asterisks, there is no statistical difference.

[a]The test in reading cannot be compared to 2016 and therefore 2017 is not included in the survey.

Effects for Group 1 in Ninth Grade (Analysis 1)

A positive number in the Effect column on the right would indicate that Pilot 1 schools have developed more positively compared to other schools in Bærum after the introduction of tablets in 2017; however, the numbers are negative. The analysis was carried out through a survey in 2017. Furthermore, the result of the previous report (Krumsvik, Berrum & Jones, 2018) is included in the first grey column to compare the short-term effect (2015, 2016) against more long-term effects (2017).

The analysis was completed in the ninth grade, as students in the eighth grade may have attended one of the primary schools that had already introduced tablets, thus creating uncertainty about the results.

National tests in reading cannot be compared to 2016, as changes have been made to the scale of this test (cf. last report, Krumsvik, Berrum & Jones 2018). Therefore, the result for reading is omitted from the analysis, as the measurement takes place in 2017.

None of the results are statistically significant, and therefore we cannot say with certainty that the negative difference is not random. This applies to both the results from 2015/2016 and 2017. However, the same trend with negative results that appeared in 2015/2016 (short term) continued in a slightly longer term time frame. in 2017.

Effects for Group 2 in Ninth Grade (Analysis 1)

Table 6 shows the average test results in the national test in arithmetic and reading for all children, boys and girls, in ninth grade for Pilot 1 schools and other schools. A positive number in the Effect column on the right would suggest that Pilot 2 schools have developed more positively than other schools after the introduction of tablets; however, the numbers are negative. Both 2016 and 2017 are included in the aftermath, which means that the measured effect is an average of the effect in 2016 and 2017.

Table 6
Difference-in-difference Analysis of Average Test Results, Ninth Grade (Group 2)

		Before tablet (2014,2015)		After tablet (2016,2017)		Effect (2016,2017)
		Pilot schools	Non pilot schools	Pilot schools	Non pilot schools	**Diff-in-diff**
All	Reading [a)]	58,6	56,8	59,3	58,1	-0,6
	Arithmetic	58,7	57,1	58,3	57,6	-0,9
Boys	Reading [a)]	58,0	55,6	58,7	56,7	-0,4
	Arithmetic	59,8	58,0	60,0	58,5	-0,4
Girls	Reading [a)]	59,0	57,9	60,0	60,2	-1,2
	Arithmetic	57,3	56,2	56,7	56,5	-1,0
Number of schools [b)]		**3 schools**	**7 schools**	**3 schools**	**7 schools**	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.

[a]The test in reading cannot be compared to 2016 and therefore 2017 is not included in the survey. [b]The number of schools is not equal to the number of observations. There are double numbers of observations in both the pre-measurements and the post-measurements, as both measurements extend over two years, i.e. all schools are included twice. The test in reading is, however, an exception, as only 2016 is included in the reassessment.

None of the results are statistically significant and therefore we cannot say with certainty that the difference is not random.

Effects for Group 1 in Ninth Grade (Analysis 2)

The fixed effect analysis in Table 7 shows the same results as the difference-in-difference analysis in Table 6. This can be seen in the variable "Effect of tablet" in Table 7 where the effect is not significant, which in turn means that we cannot conclude with certainty that there is a difference in the development of pilot schools (Group 1) as compared with other schools.

Table 7
Difference-in-difference in Number of Pupils per School Year. Fixed Effect Regression Analysis in Ninth Grade (Group 1)

	All pupils	Boys	Girls
	Arithmetic	Arithmetic	Arithmetic
Group school	1	-1,8	1,1
After implementation	1,4	0,5	-0,3
Effect of tablet [a)]	**-2,7**	**-0,7**	**-2,3**
Big school [b)]	-6,5**	1,7	-1,3
Number of pupils per year	-0,2	0,5	1,3
FE, school level	Ja	Ja	Ja
Explanatory power (R^2)	0,84	0,73	0,43
Number of observations	18	18	19

Note: Significance tests have been conducted with a linear regression analysis with fixed effect at school and year. If a number does not include asterisks, there is no statistical difference.

[a]The effect of tablets is an interaction between a dummy variable to be the input school and dummy variable to be after the implementation of the bet, i.e., the effect is calculated by a difference-in-difference approach. [b]A big school is defined as a school with 400 students or more.

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

However, we find a significant result for the fixed effect analysis, as it turns out that students in a "Big school" with more than 400 students have significantly lower test results in the ninth grade than do students in smaller schools.

The analysis is performed for 2017 and reading is therefore excluded from the analysis, cf. reasoned justifications (Berrum, Paaske Gulbrandsen, Fyhn Elgaard & Krumsvik (2018).

Effects for Group 2 in Ninth Grade (Analysis 2)

The fixed effect analysis in Table 8 shows the same results as the difference-in-difference analysis in Table 6. This can be read from the variable "Effect of tablet" where the effects are not significant and it cannot be concluded that there is a difference in the development of pilot schools (Group 2) as compared with other schools.

Table 8
Diff-in-diff in Fixed Effect Regression Analysis in Ninth Grade (Group 2)

	All pupils		Boys		Girls	
	Aritmetic	Reading [c)]	Arithmetic	Reading [c)]	Arithmetic	Reading [c)]
Group schools	1,6	0	0,4	2	3,0	2,8
After implementation	0,7	1,3**	0,1	1*	-0,1	2*
Effect of tablet [a)]	**-1,4**	**-0,5**	**-0,2**	**-0,5**	**-1,2**	**-1,0**
Big school [b)]	1,2	2,9***	2	4**	-1,6	1,3
Number of pupils per year	0,7	-0,4	0,6	0,4	1,4*	0,1
FE, school level	Ja	Ja	Ja	Ja	Ja	Ja
Explanatory power (R^2)	0,57	0,79	0,55	0,90	0,47	0,59
Number of obs.	40	30	38	29	39	28

Note: Significance tests have been conducted with a linear regression analysis with fixed effect at school and year. If a number does not include asterisks, there is no statistical difference.

[a]The effect of tablets is an interaction between a dummy variable to be the input school and dummy variable to be after the implementation of the bet. I.e. The effect is calculated by a difference-in-difference approach. [b]A big school is defined as a school with 400 students or more. [c]The test in reading cannot be compared to 2016 and therefore 2017 is not included in the survey.

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

4.1.7. Results from National Mapping Tests in First to Third Grades

In the national mapping tests, it is examined whether the students are above or below the concern threshold for the expected learning level. An increase in the proportion of students across the critical boundary at pilot schools may indicate that the introduction of tablets has contributed to increased learning from the first to third grades.

Effects for Group 1 in First Through Third Grade

Table 9 shows the proportion of students over the critical limit in the state assessment tests for reading, where we have selected subtests spelling, reading words, and reading comprehension among several subtests, and state

assessment tests on behalf of Pilot 1 schools and other schools in Bærum. A positive value in the column "diff-in-diff" indicates a positive effect of introducing tablets.

Table 9
Difference-in-difference Analysis of Share of Students above Critical Limit: First, Second, and Third Grades (Group 1)

		Before tablet (2014)		After tablet (2017)		Effect (2017)	Effect (2015,2016)
		Pilot Schools	Non-pilot schools	Pilot schools	Non-pilot schools	**Diff-in-diff**	**Diff-in-diff** [b)]
1st grade	Spelling [c)]	90,4%	90,7%	88,4%	83,0%	5,7%	3,9%
	Read words	91,7%	91,3%	91,8%	83,1%	8,3%	2,6%
	Reading com-prehension	87,6%	90,3%	90,0%	81,5%	11,1%	6,6%
	Arithmetic	80,1%	84,9%	82,3%	75,9%	11,2%	5,4%
2nd grade	Spelling [c)]	86,8%	86,0%	88,4%	83,6%	4,1%	2,0%
	Read words	89,1%	84,0%	94,4%	84,1%	5,2%	1,7%
	Reading com-prehension	81,7%	84,2%	84,3%	85,7%	1,0%	6,2%
	Arithmetic	84,3%	85,7%	85,2%	84,2%	2,4%	6,0%
3rd grade	Spelling [c)]	87,8%	88,0%	88,9%	87,3%	1,9%	0,5%
	Read words	88,2%	87,7%	88,4%	80,3%	7,6%	0,8%
	Reading com-prehension	88,7%	91,0%	88,8%	85,1%	6,0%	-0,7%
	Arithmetic	94,5%	86,5%	89,9%	87,2%	-5,4%	-15,3%**
Number of schools [a)]		**2 schools**	**18 schools**	**3 schools**	**18 schools**		

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.

[a]Bekkestua Primary School is only included in the post-measurement. Therefore, there are two schools in the pre-measurement and three schools in the post-measurement. [b]The difference is listed in percentage points. [c]In the first step, six parameters are usually measured. In this analysis, we have only used the words "Spell words" (spelling), "Read words", and "Reading comprehension". Consequently, "writing letters", "finding sounds in words", and "joining sounds" is not included in the analysis for the first grade, although this is also part of the state survey. For the second and third grades, we have omitted "Understanding words".

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

The table gives no clear conclusions. In general, effect sizes for first and second grade are positive both for 2015/2016 and 2017, but none of these can be considered to be different from zero. For the third grade, there was a significant negative effect in 2015/2016 on arithmetic. The effect is still negative in 2017, but we cannot conclude with statistical certainty that this is different from zero. This can in itself be regarded as a positive development.

Effects for Group 2 in First Through Third Grade

Table 10 shows the percentage of students above the critical boundary in the state assessment tests for reading, where we have selected the spelling, reading words, and reading comprehension among multiple subtests, and the state survey tests for pilot 2 schools and other schools in Bærum. A positive

value in the column "diff-in-diff" indicates a positive effect of introducing tablets.

Table 10
Difference-in-difference Analysis of Share of Students above Critical Limit: First, Second, and Third grades (Group 2)

	Before tablet (2014,2015)		After tablet (2016,2017)		Effect (2016,2017)
	Pilot schools	**Non-pilot schools**	**Pilot schools**	**Non-pilot schools**	**Diff-in-diff[b)]**
1st grade Spelling [c)]	90,4%	88,5%	90,8%	85,4%	3,5%
Read words	89,6%	88,8%	92,3%	85,2%	6,3%*
Reading comprehension	89,3%	88,0%	93,3%	84,8%	7,2%*
Arithmetic	85,5%	81,2%	89,4%	79,2%	5,8%
2nd grade Spelling [c)]	85,6%	86,2%	86,3%	84,6%	2,3%
Read words	86,1%	84,4%	88,3%	84,6%	1,9%
Reading comprehension	83,7%	86,9%	86,1%	84,4%	5,0%
Arithmetic	86,7%	86,2%	84,8%	83,0%	1,3%
3rd grade Spelling [c)]	88,2%	88,0%	90,2%	86,0%	4,0%
Read words	89,0%	86,4%	85,0%	80,9%	1,4%
Reading comprehension	92,7%	91,1%	92,1%	87,9%	2,6%
Arithmetic	88,2%	87,4%	87,6%	86,8%	0,0%
Number of schools [a)]	**7 schools**	**18 schools**	**7 schools**	**18 schools**	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.

[a]The number of schools is not equal to the number of observations. There are double numbers of observations in both the pre-measurements and the post-measurements, as both measurements extend over two years, i.e. all schools are included twice. The test in reading is, however, an exception, as only 2016 is included in the reassessment. [b] The difference is listed in percentage points. [c] In the first step, six parameters are usually measured. In this analysis, we have only used the words "Spell words" (spelling), "Read words", and "Reading comprehension". Consequently, "writing letters", "finding sounds in words", and "joining sounds" is not included in the analysis for the first grade, although this is also part of the state survey. For the second and third grade, we have omitted "Understanding words".

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

The table shows only positive effect sizes, but only two of the results can be considered to be different from zero. There are positive effects on reading and understanding in the first step, both of which are statistically significant.

4.1.8. Results from the Student Survey at Primary School and Secondary School

Effects for Group 1 in Seventh Grade

Table 11 shows the effect of introducing tablets in seventh grade for Pilot 1 schools compared to the other schools in Bærum. A positive value means that Pilot 1 schools have had an increase as compared to the other schools.

Table 11
Difference-in-Difference Analysis of Student Survey Indicators, Seventh Grade (Group 1)

	Effect (2015)	Effect (2016)	Before tablet (2014)		After tablet (2017)		Effect (2017)
	Diff-in-diff	Diff-in-diff	Pilot schools	Non-pilot schools	Pilot schools	Non-pilot schools	Diff-in-diff
Academic challenge	-0,10	-0,10	4,30	4,20	4,25	4,28	-0,13
Learning environment	0,12	0,00	3,90	3,75	3,80	3,79	-0,14
Bullying	-0,09	0,00	1,25	1,21	1,30	1,23	0,04
Mastery	0,01	0,06	4,10	4,08	4,00	3,90	0,08
Motivation	0,01	0,23	3,60	3,63	3,65	3,48	0,20
Enjoyment	0,05	0,20	4,20	4,24	4,30	4,13	0,21
Common rules	-0,04	0,01	4,00	3,76	4,00	3,84	-0,08
Teacher support	0,15	0,16	4,00	3,93	3,95	3,86	0,01
Home support	-0,15	0,08	4,10	4,08	4,10	4,10	-0,03
Assessment for learning	0,00	0,25	3,25	3,25	3,40	3,15	0,25
Number of schools			2 schools	8 schools	2 schools	8 schools	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.

In the student survey, the students respond on a scale from 1 to 5. The 10 indicators are based on a number of sub-questions. The composition of the indicators is described in more detail at www.skoleporten.udir.no.

*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

The table shows that there are no major differences in the student survey between Pilot 1 schools and other schools. By 2015, there was a significant effect to be found in the indicator bullying, which means that Pilot 1 schools had experienced a significant increase in bullying from 2014 to 2015. The bullying indicator is still higher for pilot schools than for other schools in 2016 and 2017, but the effect is no longer significant, which means we cannot conclude that the effect of tablets on bullying is different from zero. This means that there was a negative effect of tablets in the short term, but that effect has decreased and ceased in the long run. In addition, we cannot rule out that the impact on bullying in 2015 was influenced by a possible grade effect and other conditions, which are not related to the introduction of tablets.

To investigate the bullying results more closely, we examined the question about digital bullying in the National Student Survey (The Norwegian Directorate of Education and Training, 2018) between Group 1 and Group 2 and other schools in Bærum in 2016 and 2017. We found no significant differences in level and development between these school groups – neither combined, nor between genders. This can also be an indication that identification of bullying among girls in seventh grade in Pilot 2 schools depends on variables in addition to the usage of tablets.

Effects for Group 1 in Tenth Grade

Table 12 shows the effect of introducing tablets in the tenth grade for Pilot 1 schools as compared to the other schools in Bærum. A positive value means that pilot 2 schools have had an increase as compared to other schools in the other groups.

Table 12
Difference-in-difference Analysis of Student Survey Indicators, Tenth Grade (Group 1)

	Effect (2015)	Effect (2016)	Before tablet (2014)		After tablet (2017)		Effect (2017)
	Diff-in-diff	Diff-in-diff	Pilot schools	Non-pilot schools	Pilot schools	Non-pilot schools	Diff-in-diff
Academic challenge	-0,10	-0,10	4,30	4,20	4,25	4,28	-0,13
Learning environment	0,12	0,00	3,90	3,75	3,80	3,79	-0,14
Bullying	-0,09	0,00	1,25	1,21	1,30	1,23	0,04
Mastery	0,01	0,06	4,10	4,08	4,00	3,90	0,08
Motivation	0,01	0,23	3,60	3,63	3,65	3,48	0,20
Enjoyment	0,05	0,20	4,20	4,24	4,30	4,13	0,21
Common rules	-0,04	0,01	4,00	3,76	4,00	3,84	-0,08
Teacher support	0,15	0,16	4,00	3,93	3,95	3,86	0,01
Home support	-0,15	0,08	4,10	4,08	4,10	4,10	-0,03
Assessment for learning	0,00	0,25	3,25	3,25	3,40	3,15	0,25
Number of schools			2 schools	8 schools	2 schools	8 schools	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.

In the student survey, the students respond on a scale from 1 to 5. The 10 indicators are based on a number of sub-questions. The composition of the indicators is described in more detail at www.skoleporten.udir.no.

The table shows, as in the seventh grade (Group 1), that the effects of introducing tablets on the student's well-being and learning environment are close to zero and not significant.

Effects for Group 2 in Tenth Grade

Table 13 shows the effect of introducing tablets in the tenth grade for pilot 2 schools as compared to other schools in Bærum. A positive value means that the pilot 2 schools have had an increase as compared to the other schools. The table shows, unlike in the seventh grade in pilot 2 schools, that the introduction of tablets has not had a negative impact on the bullying indicator for tenth-grade students. At the same time, we register positive significant effects on mastering, motivation, well-being, teacher support, and assessment of learning.

Table 13
Difference-in-difference Analysis of Student Survey Indicators, Tenth Grade (Group 2)

	Before tablets (2014,2015)		After tablets (2016,2017)		Effect
	Pilot schools	Non-pilot schools	Pilot schools	Non pilot schools	Diff-in-diff
Academic challenge	4,32	4,26	4,42	4,29	0,08
Learning environment	3,68	3,79	3,97	3,82	0,25
Bullying	1,20	1,21	1,22	1,22	0,00
Mastery	4,00	4,07	4,03	3,96	0,15*
Motivation	3,53	3,61	3,62	3,49	0,2**
Enjoyment	4,17	4,27	4,33	4,18	0,25**
Common rules	3,75	3,83	3,93	3,84	0,17
Teacher support	3,87	3,94	4,03	3,89	0,22*
Home support	4,03	4,11	4,17	4,10	0,14
Assessment for learning	3,17	3,26	3,42	3,15	0,36***
Number of schools	3 schools	8 schools	3 schools	8 schools	

Note: The significance is tested by a two-tailed independent T-test with equal variance of 10%, 5% and 1% significance level. If a number does not include asterisks, there is no statistical difference.
In the student survey, the students respond on a scale from 1 to 5. The 10 indicators are based on a number of sub-questions. The composition of the indicators is described in more detail at www.skoleporten.udir.no.
*with 90 % certainty there is a difference between the effort group and the control group. **with 95 % certainty there is a difference between the effort group and the control group. ***with 99 % certainty there is a difference between the effort group and the control group.

4.1.9. On the Use of Data from National Tests, National Mapping Survey, and the National Student Survey

We repeat that it is important to note that the effect results from national tests, the national mapping survey, and the National Student Survey (The Norwegian Directorate of Education and Training, 2018) belong to different students in the pre-and post-measurements. This means that the results from the effect measurements may potentially be the result of possible grade effects. Analyses and further investigation of the results of national tests in the eighth and ninth grades showed that the results here were quite robust in regard to the grade effect, while the analysis of the state mapping tests showed that the results here were not robust in regard to the grade effect. Therefore, we cannot rule out that the grade effect may also have an impact on the results in the student survey in the seventh and tenth grades. We therefore request that the results be read with this reservation.

5. Discussion

The context for this study has been the implementation of tablets as a part of the school development in the Bærum Municipality. As Fullan (2001, 2013) mentions, it can be a challenge to carry out school leadership in a culture of change, and the study has revealed several obstacles in this implementation process of tablets in school (this is described more thoroughly in the main report by Berrum, et al. 2018).

The study shows that in several school areas, tablets have a rather limited effect on pupils' learning outcomes. It is important to underline that the study does not find any direct causality in the relationship between implementing tablets and positive learning outcomes.

However, among the significant findings in this study, we see that tablets have somewhat more positive effects among boys than among girls. The positive effect of tablets that we see among boys can be related to the fact that the use of tablets serves as a positive structuring factor for the boys' learning work. We also find support for this in the 10th grade, where boys who make use of tablets to a significantly greater extent experience having common rules for the teaching than boys in schools that do not use tablets. This may be because use of the tablet requires structure (we also find support for this in the qualitative interviews in the study). One possible explanation here can be that teachers make greater use of and make available work schedules and learning resources for school hours with the use of tablets. At the same time, the use of tablets contributes to the pupils having most of their tools and previous learning work gathered in one place in the tablet. This means that the pupils can get started quickly, and that they experience the learning resources as more transparent and accessible. We also find support for this in the qualitative data in the study.

Furthermore, it seems that the tablet can be a motivating factor in the pupils' school life. In this regard, we see significant positive findings in the 10th grade, generally for increased motivation. It seems here that the tablet device helps to make boys more motivated for learning with the use of tablets. It can also be that the tablet's multiple digital, graphic, auditory and visual capabilities and support features (visualization, audio, multimodal aspect, communication capabilities) can give new opportunities for adapted education and differentiation. There is also the possibility that the tablet device provides the opportunity for a digital support that particularly benefits low-performing students, where boys are over-represented.

Does the tablet have an equalizing effect between the sexes? And can the use of tablets in schools thus contribute to a school with less difference between girls' and boys' school performance? Today, girls generally perform better than boys, and several studies reveal that there is not any "quick fix" for increasing boys' school performance with or without educational technology. However, findings from the study suggest the possibility that boys benefit from tablets to a greater extent than girls. An interesting finding is that the effect of introducing tablets is significantly positive for boys in fifth grade in English (as in 2015/2016). Furthermore, the effect is also positive and significant for all children in fifth grade in English in 2017. These findings can be based on a number of explanations (e.g., the gaming culture among boys, etc) where tablets might only be one of several factors. In general, the study shows that the large schools especially have positive results.

From a critical point of view, one might ask if this extensive use of digital tools both in school and outside school affect pupils writing skills with pen

and paper (van der Meer & van der Weel, 2017). The study has not examined this area, but it is important for future research to raise the awareness around such digital "pitfalls".

6. Conclusion

It is still too early to say anything definitive about the effect of tablets on learning outcomes, as changes take time. We refer therefore to the effects we see in the Pilot 1 schools as "intermediate effects".

The preliminary results give reason to assume that in several subject areas, tablets have a rather limited effect on pupils' learning outcomes. However, the use of tablets can have some small positive effects on boys' learning. This can be linked to the fact that the tablet provides poorly performing students, where boys are over-represented, a digital support that contributes to smoothing the students' performance. This also presupposes an appropriate use of tablets and good teaching quality (in line with Genlott & Grönlund, 2016). The use of the tablet is strongly linked to pedagogical practice, which in turn is influenced by teacher competence. This might also link to "outside school learning" where the significantly positive results for boys in fifth grade in English can be interpreted as "a sign of the times" where English language immersion in leisure time among boys is continuously developing.

From the study, we find some tendencies that when the use of tablets is supported by teachers who have digital competence, their use seems to have a small equalising effect between the school achievements of boys and girls. However, we cannot rule out that a grade effect may also have an impact on the results, and we therefore request that the results be read with this reservation.

7. Limitations

There are some limitations in this study. First, in this part of the trailing research, we have only presented quantitative data. This might be a certain limitation since the research consists of several other data sources which give a broader picture of the implementation of tablets in Bærum Municipality.

Acknowledgements

We would like to thank Johanne Fyhn and Øystein Lorvik Nilsen for their contributions to this study. We would also like to thank Bærum Municipality and Rambøll Management for giving us the opportunity to carry out trailing research on the "Everyday Digital Schooling" pilot project. Finally, we would like to thank the University of Bergen and the project "Digital competence in higher education" for their support.

References

Almås, A.G. & Krumsvik, R. (2008). Teaching in Technology-Rich Classrooms: is there a gap between teachers' intentions and ICT

practices? *Research in Comparative and International Education 2*(3), 103–121. https://doi.org/10.2304/rcie.2008.3.2.103

Berrum, E., Paaske Gulbrandsen, I., Fyhn Elgaard, J. & Krumsvik, R.J. (2018). *Sluttrapport for Evaluering av «Digital skolehverdag»*, del II. Oslo: Rambøll Management.

Bærum Municipality (2015). *Everyday Digital Schooling – review and analysis of the experiences acquired by five pilot schools in the Municipality of Bærum*. Oslo: Bærum Municipality.

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

Cheung, A.C.K. & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review 9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

Cole, M. (1996). *Cultural psychology. A once and a future discipline*. Cambridge, MA: Harvard University Press.

Directorate of Education? If this is the "author" of a source, add it here (or, Norwegian Directorate of Education?)

Escueta, M., Quan, V., Joshua, A. & Oreopoulos, N.P. (2017). Education Technology: An evidence-based review. *NBER working paper series*. National Bureau of Economic Research. Working Paper 2374, https://doi.org/10.3102/000283120380048

Farlie, R. W. & Robinson, J. (2013). Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren. *American Economic Journal: Applied Economics*, *3*(5), 211–240. https://doi.org/10.3386/w19060

Finne, H., Levin, M. & Nilssen, T. (1995). Trailing research. A model for useful program evaluation. *Evaluation*, *1*, 11–31. https://doi.org/10.1177/135638909500100102

Fetters, M., Curry, L.A. & Creswell, J.W. (2013). Achieving Integration in Mixed Methods Designs—Principles and Practices. *Health Research and Educational Trust, 6*(48), *2134–2156.*

Fullan, M. (2001). *Leading in a Culture of Change*. Retrieved 16.04.2017 from: https://doi.org/10.1111/1475-6773.12117

Fullan, N. (2013). *Great to Excellent: Launching the Next Stage of Ontario's Education Agenda*. Retrieved 20.03.2018 from: http://www.michaelful-lan.ca/wp-content/uploads/2013/09/13_Fullan_Great-to-Excellent.pdf

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

Jahnke, I., Bergström, P. Mårell-Olsson, E., Häll, L. & Kumar, S. (2017). Digital Didactical Designs as research framework: iPad integration in Nordic schools. *Computers & Education 113, 1–15.* https://doi.org/10.1016/j.compedu.2017.05.006

Krumsvik, R. (2008). Situated learning and digital competence. *Education and Information Technology, 4(*13*),* 279–290. https://doi.org/10.1007/s10639-008-9069-5

Krumsvik, R. (2009a). Ein ny digital didaktikk. I H. Otnes (red)., *Å være digital i alle fag*. Oslo: Universitetsforlaget.

Krumsvik, R. (2009b). Situated learning in the network society and digitized school. *European Journal of Teacher Education, 2*(32), 167–185. https://doi.org/10.1080/02619760802457224

Krumsvik, R. (2012). The Digital school and Teacher Education in Norway. In, *Jahrbuch Medienpädagogik 9* (pp. 455–480). Dortmund: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-94219-3_20

Krumsvik, R. & Almås, A.G. (2008). The Digital Didactic. In, R. Krumsvik (ed.), *Learning in the Network Society and Digitized School*. New York: Nova Science Publishers.

Krumsvik, R., Egelandsdal, K. Sarastuen, N.K., Jones, L. & Eikeland, O.J. (2013). *The SMIL-study. The relationship between ICT and learning outcome in upper secondary school.* UiB/KS.

Krumsvik, R.J., Berrum, E. & Jones, L. (2018). Everyday Digital Schooling – implementing tablets in Norwegian primary school. Examining outcome measures in the 1st cohort. *Nordic Journal of Digital Literacy 3* (12), 152-176.

Kongsgården, P. (2019). *Vurderingspraksiser i teknologirike læringsmiljøer*. Doctoral thesis. Bergen: University of Bergen.

Kongsgården, P. & Krumsvik, R.J. (2016). Use of tablets in primary and secondary school - a case study. *Nordic Journal of Digital Literacy*, *4*(11), 248-270.

Kongsgården, P. & Krumsvik, R. J. (2019, in press)**.** Vurdering for læring i teknologirikt læringsmiljø – en kasusstudie av selvregulert læring blant elever i videregående skole. *Nordic Studies in Education.*

Kulik, C. L., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: an updated analysis. *Computers in Human Behavior*, *7*, 75–94. http://dx.doi.org/10.1016/0747-5632(91)90030-5.

Kunnskapsdepartementet (KD) (The Ministry of Education) (2006). *Læreplanverket for den 10-årige grunnopplæringen*. Oslo: Kunnskapsdepartementet.

Kunnskapsdepartementet (KD) (The Ministry of Education) (2015). *NOU 2015: 8. Fremtidens skole – Fornyelse av fag og kompetanser.* Oslo: Kunnskapsdepartementet.

Lave, J. & Wenger, E. (1991). *Situated Learning. Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Leedy, P.D. & Ormrod, J.E. (2005). *Practical Research: Planning and Design*. New Jersey: Pearson Educational International and Prentice Hall.

Mayer, R. E. (2010). Learning with technology. In H. Dumont, Istance, D., and Benavides, F. (ed.), *The Nature of Learning. Using Research to Inspire Practice*. Paris: Center for Educational Research and Innovation, OECD.

Moltudal, S., Høydal, K. & Krumsvik, R.J. (2019). *Glimpses into real-life introduction of adaptive learning technology: A design-based research approach to personalised student learning.* Manuscript submitted for publication.

van der Meer, A. & van der Weel, F. (2017). *Only three fingers write, but the whole brain works: A high-density EEG study showing advantages of drawing over typing for learning*. *Frontiers in Psychology*, *8*, 706. https://doi.org/10.5176/2251-1865_cbp17.1

Organisation for Economic Co-operation and Development [OECD] (2008). *Measuring improvements in learning outcomes: Best practices to assess the value-added of schools*. Paris: OECD.

Organisation for Economic Co-operation and Development [OECD] (2015). *Students, Computers and Learning: Making the Connection.* PISA. Brussels: OECD Publishing.

Piaget, J. (1967). *Biology and knowledge*. Paris: Gallimard.

Säljö, R. (2005). *Lärande och kulturella redskap: Om lärprocesser och det kollektiva minnet.* Stockholm: Norstedts Akademiska Förlag.

Säljö, R. (2017). Apps and learning: A sociocultural perspective. In N. Kucirkova & G. Falloon (Eds.), *Apps, technology and younger learners: International evidence for teaching* (pp. 3–13). London: Routledge.

Stahl, G. (1993). Supporting situated interpretation. *Proceedings of the Cognitive Science Society (CogSci '93)*, (pp. 965–970). Boulder, CO. Retrieved from http://www.cs.colorado.edu/gerry/publications/conferences/1990-1997/cogsci93/CogSci.html.

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 and research synthesis. *Computers and Education, 94,* 252–275.

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*, 4–28. http://dx.doi.org/10.3102/0034654310393361

Tamim, R.M., Borokhovski, E., Pickup, D., Bernard, R.M. & El Saadi, L. (2015a). *Large-Scale, Government-Supported Educational Tablet Initiatives.* British Colombia: Commonwealth of Learning.

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

The Norwegian Directorate of Education and Training (2018a). *The general part of the national curriculum*. Oslo: The National Education Directorate.

The Norwegian Directorate for Education and Training (2018b). *Skoleporten.* Retrieved 13.04.2018 from: https://skoleporten.udir.no/

Vygotsky, L. (1978). *Mind in society*. Harvard: Harvard University Press.

Wertsch, J. (1998). *Mind as action*. New York: Oxford University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. New York: Cambridge University Press.

Author Details

Rune J. Krumsvik
Rune.Johan.Krumsvik@uib.no

Erling Berrum

Lise Øen Jones

Ingrid P. Gulbrandsen

Editors' Note: The authors have identified this as a work in progress and could not shorten it for our proceedings or alter tables created in a different language. As a result, some tables are screen captures and some data appears with non-UK numbering conventions.