Carla Haelermans

Digital Tools in Education

On Usage, Effects
and the Role of the Teacher

SNS FÖRLAG

Digital Tools in Education
On Usage, Effects, and the Role of the Teacher

Carla Haelermans

Carla Haelermans

Digital Tools in Education

*On Usage, Effects
and the Role of the Teacher*

Table of Contents

Foreword

In this book the author, Carla Haelermans, Doctor of Economics at the Top Institute for Evidence Based Education Research [TIER], Maastricht University, is taking a close look at information and communications technology (ICT) in education. Which types of ICT use in education have proven to be effective? How is it possible to bridge the gap between research and practice? To answer these and other questions, Haelermans discusses the perspective of the literature, as well as several experiments ran at secondary schools in the Netherlands to draw conclusions.

We hope this study can contribute to the contemporary debates on digital tools in education, and be of value for decision makers. The views expressed are, of course, those of the author. SNS as an organization does not take a position. The mission of SNS is to initiate and present research-based analyses of issues of importance for society.

Eva Mörk, Professor of Economics at the Uppsala University, and Caroline Hall, Doctor of Economics and a researcher at the Institute for Evaluation of Labour Market and Education Policy (IFAU), scrutinized the manuscript and provided the author with many valuable suggestions on how to improve the text, and also with ideas on how to delve further into the subject. Many thanks for that. Full responsibility still rests with the author alone.

The study is part of a research project on "Future supply of skills." The project has been made possible through funding from a reference group that is following the program. This group consists of the Confederation of Swedish Enterprise, Academic Work, Anthon B Nilsen Education, The Swedish Employment Agency, University West, the Department of Finance, The KK-Foundation, KPMG, Ledarna, Lernia, NCC, The Swedish Association of Local Authorities and Regions, SALAR Stockholm County Council, Södra skogsägarna, The Swedish National Audit Office, Vattenfall, and the Swedish National Agency for Higher Vocational Education. Chairman of

the reference group is Annika Wohlström, Head of Change Management of NCC. Representatives from these entities have provided valuable suggestions and constructive criticism. Many thanks go to the members of the group.

The reference group members and the entities they represent are in no way responsible for the analysis and the conclusions in the report. This responsibility, as mentioned, rests with the author alone.

Swedish summary /
Svensk sammanfattning

Det är av stor vikt att underlätta för elever att lära sig så mycket som möjligt i skolan, givet deras förutsättningar. Frågan är hur vi kan försäkra oss om detta, och vilken roll informations- och kommunikationsteknik (IKT) har i detta avseende. Den här rapporten visar att IKT i undervisning kan vara effektiv under vissa förutsättningar och att läraren spelar en betydelsefull roll i sammanhanget. IKT i undervisning syftar vanligen på allt som har med teknik i utbildningen att göra, däribland sådana redskap som datorer, surfplattor, smarta mobiltelefoner och interaktiva anslagstavlor men också mjukvara som undervisningsspel och digitala inlärningsverktyg samt alla utbildningsappar som finns på internet.

Skolor börjar använda teknik i undervisningen av många olika anledningar och numera investerar de ofta i digitala verktyg. Några av de vanligaste skälen är möjligheterna att individualisera barnens inlärning och maximera varje barns potential, att ge barnen individuell återkoppling, att få insyn i barnens utveckling för att utforma en arbetsplan för klassen samt att fördela resurser till barn med särskilda behov.

Hur effektiv IKT är i undervisningen beror främst på hur den tillämpas liksom på vilken typ av inlärning som den används för. Men eftersom en effektiv användning av IKT förutsätter att digitala verktyg finns på plats, är detta det första steget. Dessutom är det viktigt att skolledare och lärare ser nyttan av IKT och är trygga i att använda den. Bara det faktum att man har tillgång till IKT innebär inte nödvändigtvis att den används effektivt. Det är också viktigt att se IKT som ett redskap och inte som ett mål i sig. Dessutom är det oftast så att en del av den tid och de resurser som tidigare skulle satsats på traditionell undervisning nu i stället satsas på IKT. Därför ska IKT åtminstone leda till samma resultat för att vara effektiv.

Syftet med den här rapporten är att bidra till debatten om vilken IKT-

användning som visat sig vara effektiv i undervisningen. Detta kommer att diskuteras utifrån såväl tidigare forskning som ett antal experimentstudier på högstadiet i Nederländerna. Syftet med denna rapport är också att sammanföra forskning med praktik. Det görs genom att utförligt diskutera tillämpningen av det som framkommit i tidigare studier och de holländska experimenten.

Generellt verkar en effektiv användning av IKT i undervisningen handla om följande:

1. att öka kunskapen om vilka IKT-verktyg för undervisning som är effektiva
2. att skapa en positiv inställning till IKT-verktygen bland dem som faktiskt använder dessa, helst nerifrån och uppåt
3. att underlätta för skolor, skolledare och, viktigast av allt, lärare att bli hemmastadda med IKT-verktygen och känna sig trygga med att de använder dem på det mest effektiva sättet.

Det är värt att notera att det inte verkar vara nödvändigt med en dator per elev för att uppnå de positiva effekter som nämns i litteraturen och i de nederländska experimenten. Teorier om förändringsledning och föreslagna tillvägagångssätt kan spela en viktig roll för punkterna två och tre.

Slutsatser utifrån forskningslitteraturen

De viktigaste slutsatserna i den internationella litteraturen i nationalekonomi angående effekterna av IKT i undervisningen är följande:

- Allmänna investeringar i IKT i utbildningen, utan specifika syften för vad man investerar i eller hur IKT ska användas i undervisningen, ger i bästa fall blandade resultat.
- Undersökningar av effekterna av datorstödd undervisning kontra traditionell undervisning i klassrum, där IKT är ett komplement för läraren, har visat på positiva effekter, men de är mycket små.
- Positiva effekter för specifika digitala undervisningsverktyg i studier genomförda i utvecklingsländer, både inom ämnena matematik och språk. För länder i västvärlden återfinns positiva effekter inom matematik men inte inom språk.

- De tre slutsatserna ovan kan sammanfattas som att effektiviteten av IKT i undervisning i hög utsträckning beror på hur den används och vilket (pedagogiskt) syfte den ska tjäna.
- Kostnadseffektiviteten hos IKT(redskap) inom utbildning undersöks sällan, så ytterligare forskning behövs för att kunna dra tillförlitliga slutsatser om detta. Men de få undersökningar som gjorts på området drar slutsatsen att IKT(redskap) är kostnadseffektiva och kostar lika mycket eller mindre än att minska klassernas storlek eller att anställa fler lärare.
- Det finns många hinder för lärare när det kommer till teknisk utveckling, vilket kan förklara varför en effektiv teknisk anpassning i skolor fortfarande inte har levt upp till förväntningarna. I vissa fall motsätter sig lärare tekniska förändringar i allmänhet, exempelvis på grund av en inre övertygelse. I andra fall vet lärare inte hur tekniken kan anpassas effektivt i klassen, vilket kan förklaras av sådant som brist på tid, kunskap eller övning.
- Det närbesläktade litteraturområdet förändringsledning kan ge en del goda förslag på hur förändringar kan genomföras i offentlig sektor, som skolor. Sådana exempel innehåller en genomförandeplan, resurstillförsel och säkerställande av stöd från högsta ledningen.

Slutsatser från försöksverksamhet i Nederländerna

De viktigaste slutsatserna från åtta olika, slumpvis sammansatta försöksverksamheter med IKT på högstadiet i Nederländerna, som beskrivs i denna rapport, är:

- Positiva effekter inom matematik och delar av språkinlärning: Den övergripande slutsatsen som dras utifrån dessa försök med IKT i undervisningen i Nederländerna är att medelpositiva och statistiskt säkra effekter visade sig inom matematik och inom vissa moment i språkinlärning. Resultaten för matematik liknar de som omnämns i den övriga forskningslitteraturen medan resultaten för språkinlärning skiljer sig från litteraturen, där man inte har funnit några effekter i utvecklade länder.

- Individualisering är effektivt: Försöken visar att man kan dra slutsatsen att det snarast är individualiseringen av uppgifterna som gör digitala arbetsredskap effektiva och inte endast den extra övningstiden i sig. Men elever som övar mer erfar även mer statistiskt säkra effekter.
- Effektivt för uppgifter som är enkla att automatisera: Resultaten tyder på att digitala verktyg som används för att anpassa uppgifterna till varje elevs nivå är mycket lovande för att öva upp grundläggande färdigheter inom matematik, språk och biologi. Exempelvis som addition och multiplikation inom matematik och stavning inom språk. Viktiga förutsättningar är att uppgifterna lätt kan automatiseras och att eleverna stimuleras att använda de digitala läromedlen fullt ut.
- Effekterna varierar mellan områden: Resultaten av studierna visar att det finns skillnader i hur stor nytta eleverna har av digitala inlärningsverktyg beroende bland annat på elevens ålder och inom vilket område verktyget används. Rent generellt är de enklare områdena inom matematik, stavning och grammatik mest givande för elever i årskurs 7 och 8 medan de svårare områdena inom matematik är mer givande för högpresterande elever och elever i årskurs 8 och 9.
- Effekterna varierar mellan olika prestationsnivåer: Det finns statistiskt säkra skillnader på effekten hos anpassade digitala läromedel mellan låg-, medel- och högpresterande elevgrupper. Även om det konstaterats att både låg- och medelpresterande elevers allmänna matematiska färdigheter ökar betydligt när de övar med hjälp av nätbaserade redskap, så är effekten betydligt mindre för de medelpresterande än för de lågpresterande. Det är oklart om det finns någon allmän effekt för de högpresterande (även om det, som tidigare nämnts, finns positiva effekter inom de svårare matematiska områdena). Skillnaden i effekt mellan låg- och medelpresterande elever märks även i analyserna av de olika matematiska färdighetsområdena (tal, proportionalitet och mätning). Lågpresterande elever har avsevärd nytta av det nya inslaget i undervisningen på samtliga matematikområden medan medelpresterande elever har betydande nytta inom områdena proportionalitet och mätning.
- Indelning av klassen i olika grupper med hjälp av hårdvara är effektivt: Man har funnit statistiskt säkra och positiva effekter av hårdvaruanvändning, exempelvis interaktiva anslagstavlor, för högstadi-

ets lägre årsklasser. Dessa effekter verkar dock främst bero på den differentiering som blev möjlig genom användandet av interaktiva anslagstavlor. Dessutom visade denna studie att det var avgörande att läraren fått adekvat utbildning i förväg.

- Digitala prov är effektiva: Man har funnit positiva effekter av att använda digitala prov som inlärningsredskap, vilket är snarlikt det som framkommer i litteraturen. Effekterna förklaras ofta bero på ökad exponering för materialet och/eller upprepningsprocessen som återaktiverar minnet. Ett försök visar att enkla digitala prov, även sådana utan omfattande återkoppling, ger positiva resultat.

- Digital återkoppling är effektiv: Som tillägg till föregående punkt kan nämnas att ännu större positiva effekter noterats om digital återkoppling ges vid digitala prov. Undervisningsresultaten förbättras när lärare inkluderar flervalsfrågor med utökad personlig återkoppling i sina klasser. Resultaten är signifikant bättre än bland elever som fått samma veckovisa prov men nästan helt utan återkoppling. Effekten är ungefär densamma i samtliga grupper som deltagit.

- Effekterna beror ofta på läraren: Hur effektiv IKT i undervisningen är i många fall relaterat till det sätt som läraren implementerar IKT-förnyelsen och vilken kunskap läraren har om hur IKT kan användas, vilket gör lärarens roll viktig för hur effektiv IKT är i undervisningen.

- Föräldrars engagemang är viktigt: Föräldrar kan spela en viktig roll för att stimulera elever i högstadiets lägre årsklasser att öva med anpassade digitala läromedel online. En statistiskt säker effekt gäller sambandet mellan föräldrars engagemang i elevens övningsbeteende, och därmed också elevens prestationer i matematik i årskurs 7 och 8. Effekterna av föräldrarnas engagemang är särskilt tydliga för elever i högstadiets lägre årskurser.

- Sammanfattningsvis: Experimentstudierna i Nederländerna visar att syftet med, och på vilket sätt IKT används i undervisningen, har betydelse för effektiviteten av IKT i utbildning.

Rekommendationer

Baserat på ovanstående kan rekommendationerna delas in i tre kategorier: rekommendationer på skolnivå, rekommendationer på nationell nivå och rekommendationer på allmän nivå.

Skolnivå

- Övervägt val och införande av IKT-redskap: Skolor bör försäkra sig om, och underlätta för att information samlas in om nya IKT-redskap samt under vilka omständigheter de kan tänkas vara effektiva. Detta baserat på vetenskaplig forskning. Skolor bör också försäkra sig om vilket genomförande och vilket digitalt läromedel som behövs för att det ska bli så effektivt som möjligt. Skolor bör också tillåta, eller rent av försäkra sig om, att forskning genomförs samtidigt som nya IKT-redskap utprövas. Detta för att få kunskap om effektiviteten hos ett specifikt digitalt läromedel i en svensk kontext.
- Fortbildning för yrkesmässig utveckling: Skolhuvudmän och skolledare bör lägga större vikt vid fortbildning för lärare i allmänhet och skapa utvecklingsmöjligheter för samtliga lärare i de fall detta inte är möjligt. Det utvecklar lärarnas färdigheter och tänkesätt ytterligare, så att de kan välja de bästa pedagogiska metoderna i sin undervisning.
- Underifrån och upp-metod: Skolor bör stödja en entusiastisk lärare med innovativa idéer om hur man kan tillämpa IKT på ett effektivt sätt i klassrummet. Hur IKT kan användas som ett inlärningsredskap i stället för ett administrationsredskap. En välinformerad, entusiastisk och engagerad lärare kommer att sprida informationen vidare till kollegor som undervisar i samma ämne, som i sin tur sprider det vidare till kollegor som undervisar i andra ämnen, vilket slutligen leder till att nästan hela skolan är involverad. Sedan sprids det till en annan skola och så vidare.

Nationell nivå

- Mer kunskap om effekter av digitala läromedel: Mer vetenskaplig forskning behövs för att studera kausala effekter av olika användning av IKT i klassrummet, i Sverige och i andra nordiska länder med liknande undervisningssystem. För närvarande finns bara en begränsad mängd belägg tillgängliga i fråga om vad som fungerar och vad som inte fungerar avseende IKT i dessa länder.

- Kunskapssystem och infrastruktur på nationell nivå: Regering och myndigheter i Sverige spelar en viktig roll, inte bara för att entusiasmera de verksamma inom utbildningsområdet för IKT i klassrummet, utan också för att ta fram olika effektiva användningssätt och delta i forskning om dessa frågor. Staten bör se till att kunskap sprids om att det inte bara gäller att inneha IKT (utrustning) utan också hur IKT kan användas på ett effektivt sätt i undervisningen och vilka bevisat effektiva valmöjligheter som finns. Det är särskilt viktigt eftersom inte alla sätt att använda IKT är effektiva. Den svenska staten kan också dra lärdomar från Nederländerna, när det gäller behovet av en nationell organisation för att nå de pedagogiska utövarna och informera dem om vetenskapliga rön inom området. Men också att stimulera forskning om effekterna av IKT i undervisningen och/eller erbjuda mer vägledning om effektiv användning av IKT i undervisningen. En möjlighet skulle vara att göra det till ett av Skolverkets uppdrag.

- Sprida forskningsresultat: Staten har ansvar för att göra resultaten av vetenskaplig forskning nåbara på ett tillgängligt språk för alla intressenter inom skolans område. Eller med andra ord se till att forskningsresultat om effektiv användning av IKT i undervisningen når dem som fattar beslut om hur IKT ska användas.

- Integrera IKT i lärarutbildningen: Staten har också möjlighet att se till att elevernas lärare lär sig mer om olika sätt att använda IKT i den dagliga undervisningen. Detta genom att göra effektiv IKT-användning till en del av lärarutbildningen inom högskolan. De som redan är lärare bör också stimuleras och få möjlighet att delta i dessa kurser som en del i deras livslånga lärande.

- Integrera IKT i utbildningsplaner: Staten kan även stimulera lärare att åtminstone tänka kritiskt på hur de kan använda IKT på ett effektivt sätt i sina klasser och fatta välavvägda beslut om IKT-

användning genom att ta med tänkbara sådana användningsområden i samband med att de utformar undervisningsplaner för olika klasser.

Allmän nivå

- Fokus på effektiv tillämpning: Skolhuvudmän, skolledare och myndigheter kan, och bör, stimulera lärare att börja använda IKT på ett effektivt sätt. Samtidigt bör forskning tillåtas att genomföras parallellt för att lära mer om effektiviteten av olika specifika digitala läromedel i Sverige. Detta skulle medföra att man inte bara fokuserar på att använda IKT, utan mer specifikt på hur IKT kan användas effektivt i undervisningen, exempelvis genom att undersöka hur IKT kan hjälpa lärare att uppnå sina mål mer ändamålsenligt.

- Underskatta inte betydelsen av den mänskliga faktorn: En viktig, men ofta förbisedd/undervärderad, aspekt av IKT i undervisning är den mänskliga faktorn: lärares och skolledares uppfattningar och attityder till IKT (och kanske till förändring i allmänhet). Att införa IKT handlar inte bara om att ha redskapen och att ge lärarna rätt utbildning i hur digitala hjälpmedel fungerar och kan användas. Det handlar också till stor del om huruvida lärarna tror att de digitala verktygen kommer att förbättra undervisningen eller inte. Att lärarna är skeptiska och inte ser nyttan med IKT kan vara ett stort hinder. Ett framgångsrikt införande av digitala inlärningsverktyg handlar därför också mycket om ledarskap.

English Summary

It is of great importance to ensure that students perform at the highest possible level, using their innate abilities. However, the question is how we can ensure that, and what role information and communications technology (ICT) can play in this. This report shows that ICT in education can be effective under certain circumstances, and that the teacher plays a significant role in this. ICT in education generally refers to anything that involves technology in education. This includes devices such as computers, tablets, smartphones and interactive whiteboards, but also software such as educational games and digital learning tools and all educational applications that can be found on the Internet.

Schools start using technology in education for a variety of reasons, and nowadays they often invest in digital tools. The potential to individualize children's learning and maximize each child's potential, giving individual feedback to children, gaining insights into children's progress to develop a work plan for the class and allocating resources to children with extra needs are some of the most common purposes of introducing digital tools in education.

The effectiveness of ICT in education is primarily dependent upon the way ICT is implemented, as well as on the types of learning for which ICT in education is used. However, since ICT can only be effective if there is ICT at all, this is the first step. Furthermore, it is important that school leaders and teachers see the benefits of ICT and are confident in using it. Simply having access to ICT in education will not necessarily lead to its effective use, and might even lead to negative results if ICT is merely a distraction and not applied in an effective way. It is also important to see ICT as a tool, and not as a goal in itself. Furthermore, in most cases, some of the time and resources that before would be invested in traditional teaching are now invested in ICT. Therefore, overall ICT should give at least the same results in order to be effective.

This report aims to contribute to the debate on which types of ICT use in education have proven to be effective. This will be discussed from the perspective of earlier research, as well as from several studies about experiments at secondary schools in the Netherlands. Furthermore, the aim of this report is to bring research and practice closer together, by also explicitly discussing the applicability of the findings from earlier studies and the Dutch experiments.

In general, making ICT use in education more effective seems to be a matter of: 1) increasing knowledge regarding which ICT applications in education are effective, 2) creating common support for using ICT in education among those that actually have to use the ICT tool, preferably from the bottom up, and 3) facilitating that schools, school managers and, most importantly, teachers, become acquainted with the ICT application and feel confident using it in the most effective way. It is worth noting that one computer per student does not seem to be necessary to achieve the positive effects found in the literature and the Dutch experiments. Change management theories and suggested approaches can play a significant role in the second and third aspects.

Conclusions from the literature

The main conclusions from the international literature in economics on the effects of ICT in education are the following:

- The general investments in ICT in education without a specific purpose on what to invest in or how to use ICT in education provide mixed results, at best.
- Studies on the effects of computer-assisted instruction versus traditional classroom learning, where ICT is a complement to the teacher, do find positive effects, though very small ones.
- For specific digital learning tools, positive effects are found in developing countries, both for mathematics and for language. For Western countries, positive effects are found only for mathematics, but not for language training.
- In sum: the effectiveness of ICT in education is highly dependent on how it is used and which (pedagogical) purpose it is intended to serve.

- Cost-effectiveness of ICT (tools) in education is rarely studied, so more research is needed to draw reliable conclusions on this topic. However, the few studies that look into this conclude that the ICT (tool) is cost-effective and similar to or cheaper than reducing class size or hiring an additional teacher.
- There are many barriers to technological change for teachers, which might explain why effective technology adaptation in schools has not, to date, lived up to expectations. Teachers are either resisting the technological change in general, due to, for instance, their internal beliefs, or do not know how to apply the technology effectively in class, due to factors such as lack of time, knowledge or training.
- The closely related literature strand on change management may provide some good suggestions on how to implement changes in public sector organizations such as schools. Examples of these are providing a plan for implementation, providing resources and ensuring top-management support.

Conclusions from experiments in the Netherlands

In addition, the main conclusions from eight different randomized experiments on ICT in secondary education in the Netherlands, described in this report, are the following:
- Positive effects for mathematics and some aspects of language: The overall conclusion drawn from these experiments on ICT in education in the Netherlands is that medium positive and statistically significant effects are found for mathematics, and for some aspects of language learning. This finding on mathematics is similar to that found in the literature, but, for language, this finding differs from the literature, where no effects on language are found in developed countries.
- Individualization is effective: From the experiments, it can be concluded that it is rather the individualization of exercises that makes digital practice tools effective, and not merely the additional practice

time in itself. However, students that practice more also experience more statistically significant effects.

- Effective for instructions that are easy to automate: The results indicate that digital tools that are used for adapting instructions to apply for each student are very promising when it comes to practicing basic skills in mathematics, language and biology, such as addition and multiplication for mathematics and spelling for language. Important conditions are that the instructions are easy to automate and that students are stimulated to use the digital tool to its full extent.

- Effects differ across domains: The results indicate that students benefit differently from using digital tools depending on the age of the student and depending on the domain the tools are used in. In general, the easier domains of mathematics, spelling and grammar are predominantly beneficial for students from grades 7 and 8, and the more difficult domains of mathematics are more beneficial for higher performing students and students in grades 8 and 9.

- Effects differ across performance levels: There are statistically significant differences in the effect of adaptive digital learning materials between low-, middle-, and high-achieving student groups. Although it is found that the overall mathematics skills of both low- and middle-achieving students significantly increase when they practice with the online tool, it is significantly less effective for middle achievers than for low achievers. For high achievers, it is unclear whether there is an overall effect (although, as noted in the previous conclusion, positive effects are found for the more difficult mathematics domains). The differences in effect between low- and middle-achieving students are also seen in the separate analyses among mathematics skills domains (numbers, proportions, and measurement). Low-achieving students benefit significantly from the intervention in all mathematics domains, while middle-achieving students benefit significantly from the intervention in the domains, proportions and measurement.

- In-class-level differentiation through use of hardware is effective: Statistically significant and positive effects are found by using hardware such as interactive whiteboards in class in lower secondary education. However, these effects primarily seem due to the differentiation that was possible because of the use of the interactive

whiteboard. Furthermore, a crucial part of this study was that teachers were properly trained beforehand.

- Digital tests are effective: Positive effects are found using digital tests as a learning tool, which is similar to that found in the literature, where the effects are often explained by the increased amount of exposure to the material and/or the retrieval process of information that reactivates the memory. One experiment shows that simply digital testing, even without extensive feedback, already produces positive outcomes.

- Digital feedback is effective: In addition to the previous point, even higher positive effects are found if digital feedback is given when testing digitally. Educational outcomes are improved when teachers incorporate formative, multiple-choice tests with extended, personalized feedback in their classes. The results are significantly higher than for students that were given similar weekly tests with barely any feedback. The effect is similar across all performance groups.

- Effects are often dependent on the teacher: The effectiveness of ICT in education is, in many cases, related to the way the teacher implements the ICT innovation and the knowledge of the teacher on how to use the ICT innovation, making the role of the teacher important in the effectiveness of ICT in education.

- Parental involvement is important: Parents can play an important role in stimulating students in lower secondary education to practice with adaptive online tools. A statistically significant effect is found for parental involvement on student practice behavior, and, in turn, on mathematics performance for students in grades 7 and 8. The effects of parental involvement are specifically present for low-SES students.

- To summarize: The Dutch experiments show that the effectiveness of ICT in education depends on how it is used and on the pedagogical purpose of the digital tool.

Recommendations

Based on the above, the recommendations can be organized into three categories: recommendations at the school level, recommendations at the national level and recommendations as the general level.

School level

- Deliberate choice and introduction of ICT tools: Schools should ensure, and facilitate, that information is gathered about new ICT tools and under what circumstances they are likely to be effective. This should be based on scientific research, and what form of implementation and application is needed in order to make it as effective as possible. Schools should also allow, or even make sure, that research simultaneously takes place, when new ICT tools are tried out, to learn about the effectiveness of this specific application in Sweden.
- Continuous training for professional development: The head of schools and school managers should give a more prominent role to teacher development in general and provide development opportunities for all teachers, if not present already, as this will further develop teachers' skills and mindsets in order to choose the best pedagogical methods for their teaching practices.
- Bottom-up approach: Schools should support an enthusiastic teacher with innovative ideas on how to effectively apply ICT in the classroom as a learning, instead of an administrative, tool. A well-informed and dedicated teacher will spread the word first to colleagues teaching the same subject, who will then spread the word to colleagues teaching different subjects, culminating in the involvement of nearly the entire school. From one school comes another, and so on.

National level

- More evidence about the effects of digital tools: More scientific research with causal research designs is needed in order to study the effects of different ICT uses in class, in Sweden or other Nordic

countries with similar education systems, as there is currently limited evidence available regarding what works and does not work with respect to ICT in these countries.

- National knowledge system/infrastructure: There is a significant role for the national government in Sweden to not only make education practitioners enthusiastic about using ICT in the classroom, but also different ways to use it effectively, and to participate in research about this topic. The government should ensure that the knowledge is disseminated that it is not only about having ICT (devices), but also about how ICT in education can be used effectively, and which proven effective choices are available, particularly because not all ways of using ICT are effective.

- National knowledge system/infrastructure: The national government in Sweden can also learn from the Netherlands in the sense that it needs national organizations that have the aim of reaching the educational practitioners and informing them about knowledge from scientific research, as well as to stimulate research on effects of ICT in education, and/or provide stronger guidance on the effective use of ICT in education, for example by making it part of the mandate of the Swedish National Agency of Education.

- Communication of research results: The national government has a responsibility for making the results from scientific research available in accessible language to all potential stakeholders in the field of education. In other words, ensuring that research results about the effective use of ICT in education reach those who will be required to make decisions regarding its use.

- Integrating ICT in teacher education: The national government also has the opportunity to ensure that student teachers learn more about different ways of using ICT in their everyday educational practice, by making the effective use of ICT and different ways of using it in a pedagogical way part of teacher-training programs at the higher education institutes. The current teaching body should also be stimulated and facilitated to participate in these courses as part of life-long learning.

- Integrating ICT in educational plans: The national government can also stimulate that teachers at least think critically on how they can

use ICT in their classes in an effective way, and make well-considered decisions regarding ICT use, by making the ways in which ICT is used part of the considerations for teachers writing their educational plans for each class.

General level

- Focus on effective practice: The head of schools, school managers and national governments can, and should, stimulate teachers to start using ICT in an effective way, while again allowing for research to take place at the same time, in order to learn about the effectiveness of this specific application in Sweden. This would entail not only focusing on using ICT, but more specifically on *how* to effectively use ICT in education, for example by looking into how ICT can help teachers accomplish their goals more effectively and efficiently.
- Do not underestimate the role of the human factor: An important – and often overlooked or underestimated – aspect of ICT in education is the human factor: teachers' and school leaders' negative beliefs and attitudes towards ICT (and perhaps towards change in general). Introducing ICT is not only about having the tools and providing the teachers with the right training about how digital tools work and can be used, it is also to a large extent about whether the teachers believe that digital tools will improve the education. If the teachers are skeptical and do not see the use of ICT in class this may be an important barrier. A successful implementation of digital tools therefore also required good leadership.

Chapter 1.
Introduction

It is of great importance to ensure that students perform at the highest possible level, using their innate abilities. The scientific literature shows that higher student performance is associated with higher wages for these individuals (Ashenfelter et al., 1999), fewer health problems (Conti et al., 2010) and a lower chance of subsequently being involved in criminality (Groot & Maassen van den Brink, 2010), among others. However, one main question is how to ensure that students perform at the highest possible level, while another is what role information and communications technology (ICT) can play in achieving the highest possible performance.

The phrase *ICT in education* (often also referred to as IT in education, which to most people, and also in this report, means the same thing) generally refers to anything that involves technology in education. This includes devices and means of communication such as computers, tablets, smartphones, interactive whiteboards and broadband and Wi-Fi, but also software such as educational games and (adaptive) digital learning tools and all educational applications that can be found on the Internet. Of course, the first part of the definition (technical devices, means of communication) is a prerequisite for the second part of the definition (software and tools), although this report will show that, in itself, the device or even the tools in themselves are not sufficient and not necessarily effective, as it all depends on how they are used.

Many individuals and governments believe that ICT in education makes lifelong learning easier and contributes to individual development by enhancing digital skills and providing individuals with the 21st century skills that they might need in the future and thereby increasing employability and productivity. The latter can, in turn, contribute to higher innovation rates and potentially a higher growth of GDP. Furthermore, ICT is believed to

enhance the digital competence of students, and complement and make it easier for the teacher in the teaching process (OECD, 2015). ICT can be an important tool for teachers to identify deficits in individual student progress and help students, and for policy makers and researchers to gain greater insight into student performance. Furthermore, ICT in education is suggested as being one of the solutions to reduce the problem of teacher shortages (Van der Schaft, 2008), although there is no evidence to underline these suggestions. Aside from this, many scholars also conclude that individual differentiation with ICT is the key to higher student performance (e.g. Hattie, 2009). Traditional classroom settings only partly allow schools and teachers to differentiate their teaching between students, and this would be considerably easier with ICT.

Over the years, the presence of computers in education, as well as the development of individualized ICT tools aimed at developing student skills, has increased, and accordingly, many schools have started using computers and these tools. However, schools are introducing ICT in very diverse ways, and much of its potential effectiveness depends on the teachers' and students' efficiency in using the ICT tool. It is therefore perhaps not surprising that the literature shows very mixed results from the studies in which ICT tools or ICT in education in general are analyzed. On the other hand, it also seems that research and practice are not aligned to any great extent, and the findings from scientific studies on ICT in education also do not necessarily seem to be the basis for the decisions schools make regarding this topic, as anecdotal evidence from the Netherlands shows. It might even be that school policy is more about the presence of ICT, i.e. having ICT is the goal, instead of the way the ICT is used, i.e. ICT as the means of achieving the overall goal of higher student performance, given that the OECD concluded in 2015 that most schools seem to have sufficient amounts of ICT, but that the overall benefits are not yet visible in student performance in the PISA (Programme for International Student Assessment) 2012 (OECD, 2015). One of the most common critiques against ICT use in education is that in most cases, some of the time and resources that before would be invested in traditional teaching are now invested in ICT, whereas the effects of ICT are not unambiguous. Furthermore, computers appear to be often primarily used for 'normal computer use' (word processing, internet browsing, games, music, etc.) and less frequently for didactic applications, which could potentially explain the lack of convincing evidence on educational

performance. Lastly, an ICT device can, in itself (without effective use), prove distracting to students, and some studies have shown that students perform better if their smartphones or devices are banned in class (Beland & Murphy, 2016; Carter et al., 2016).

ICT in education dates back to 19th century, when the first primitive version of a slide projector was used. This was followed by, for example, the overhead projector in the first half of the 20th century and the calculator in the 1970s.[1] The development of ICT in education has accelerated ever since, with a quarter of the schools already using a personal computer for college and career guidance by 1986[2] and more than 97% of the classrooms in the US having one or more computers in 2009. When ICT first became more generally available, policy-makers stimulated ICT presence in education, with subsidies available in many countries. Although ICT in education is still present on the policy agenda, over the years the financial stimulation has decreased and the responsibility to increase ICT-use in education has shifted more towards the schools. Moreover, in most OECD countries, the focus has shifted from ICT presence to ICT use in education (OECD, 2015). Although it could be seen as highly desirable that most educational organizations have made the transition to effective use of ICT in education by now, this does not seem to be the case. Everyday practice in educational institutes, however, shows a different story. Although some schools in some countries have indeed made the transition very well and are effectively using ICT tools in the educational process, others have run into difficulties somewhere in the process. They often have ICT tools, but do not use them effectively. There are various potential reasons for this. One of these reasons could be resistance to change or lack of knowledge of how to use the tools, for example, among the faculty. Another potential reason can be found in technical issues, such as (lack of) a broadband internet connection. Furthermore, it can also be the case that the number of computers/ laptops/tablets in class is insufficient for optimal use.

Second, in the scientific research, most studies have focused on either of two extreme cases of the effects of ICT in education. Studies either belong to the strand of literature that has focused on general ICT use in education,

---

1. http://online.purdue.edu/ldt/learning-design-technology/resources/evolution-technology-classroom
2. https://web.csulb.edu/~murdock/histofcs.html

or, even more broadly defined, investments in ICT in education; or studies belonging to the strand of literature that focuses on very specific ICT use in education, such as digital differentiated practice programs (e.g. FastForWord, or Mousework). The former studies are very broad and often not sufficiently specific to draw conclusions on the effect of ICT in education, as they do not focus at all on *how* ICT in education is used. The latter group of studies are, in contrast, so specific to the specific tool that they are studying, that it is difficult so say anything about the effects of use of ICT in education in general. There are only a limited number of studies that have focused on the effects of the use of more general ICT tools (such as the smartboard or general digital learning materials instead of the general books), which is the type of knowledge for which schools might have more need.

Third, there seems to be a significant gap between the findings from academic research and the (use of) the ICT tools that can be found in schools. As mentioned, the choice of digital material is rarely based on evidence, as most schools and teachers make their decision based on their intuition or on what other schools are using. In part, this seems to be caused by a lack of knowledge in the educational sector with respect to the results from academic research. This might be due to the fact that academic research and publications are hard to access for people who are not in academia, and that academic research is often not written in an accessible language that reaches the broader public, such as the educational sector. It could also be due to the fact that academic research rarely provides indisputable evidence that all points in only one direction. Moreover, if the results of academic research are available to the broader public, they might not provide the information for which the sector is looking, as described in the previous paragraph.

It seems that effectively using ICT in education while not being able to effectively transfer knowledge from scientific research to practice is a problem in many countries, including both Sweden and the Netherlands. However, in the Netherlands, schools seem to use ICT more in class, and more research experiments take place in these schools to study the effectiveness of ICT use in education in practice. These research experiments are, among other elements, discussed in this report.

Aims and contributions

Keeping the various reasons for the lack of effective ICT use in educational institutes in mind, this report aims to contribute to the discussion on which types of ICT use in education are proven to be effective. The results will be drawn from the literature, as well as from several experiments conducted in secondary schools in the Netherlands. To be more specific, the aim of this report is threefold:

1. To provide an extensive overview of the current knowledge on the causal effects of ICT in education from the scientific literature and from several experiments that were conducted in the Netherlands. This report will primarily present and review (international literature on) experimental studies on the effects of ICT in education, where primary, secondary and vocational education will be covered, but *not* higher education.

2. To discuss the role of the most important players in education besides the students themselves, namely the teachers, in the effective implementation of ICT in education, based on both the literature and on the experience from the experiments in the Netherlands that are discussed in this report. The focus here lies on the crucial role of the teacher in the process of making ICT effective, when ICT is introduced and implemented in education.

3. Lastly, the aim is to try to take a first step towards bridging the gap between research and practice, by bringing all of the above together in this report, and in a conclusion and discussion in which the applicability of the findings from the literature and the experiments is discussed. Furthermore, some suggestions on how to deal with poor implementation, or no implementation at all, of proven effective ICT tools and devices are also provided.

Outline of the report

In the remainder of this report, Chapter 2 will present the findings of the previous international literature, first explaining which research designs are included in the literature search, namely only those that allow for causal inference. Causal inference implies that one can be sure that the studied effect in the outcome measure (e.g. student performance) is exclusively due to the studied intervention (ICT in this case), and not to other factors indirectly influencing the outcome measure (for example if smarter children with wealthy parents are more likely to have a computer than less-gifted children, this is a factor which needs to be controlled for, because otherwise it cannot be established if a potential effect is due to ICT or due to the fact that such children were anyway originally smarter). Following this, two strands in the international literature are discussed: Studies on the effect of accessibility of, and investments in, ICT in education in general, and studies on the effects of very specific ICT tools in education. In this chapter, two levels of education (primary and secondary) in both strands are covered as, to date, the literature has primarily focused on these two. The literature overview ends with a discussion on the barriers to technological change that teachers might experience, and concludes with a short overview of the effects of ICT in education and the role of the teacher within it.

Following this, in Chapter 3, the research context is discussed. First, the educational system in the Netherlands is briefly explained, focusing on the key features of this system and the major differences compared to other educational systems. Next, the Dutch educational ICT policy is explored, and an overview given of the actual ICT use in Dutch schools, using some summary statistics. The chapter concludes with all relevant aspects of the Dutch context compared to that of the Swedish context.

In Chapter 4, several studies on the effects of ICT on education in the Netherlands will be discussed. The chapter consists of a brief description of studies on the effects of the digital differentiated program Mousework (focusing on mathematics and literacy), the effects of GotiIT?! (focusing on mathematics performance), the effects of the use of an interactive whiteboard (SMARTboard) in class, and the effects of digital differentiation, digital feedback and digital weekly testing. For each ICT application studied, a description of the research is first provided, in which the characteristics of the program under study are outlined, along with the setup of the

experiments and a short description of the school(s) under study. Subsequently, the findings on the effects of each application are given. Chapter 4 concludes with a synopsis in which the findings of each program are amalgamated into general findings on the effects of using ICT in education in secondary schools in the Netherlands.

Chapter 5 consists of the concluding remarks, starting with the general conclusions of the report, followed by the discussion of the applicability of the findings for educational practices in both the Netherlands and Sweden, along with some policy recommendations.

Chapter 2.
Literature Review,
Effects of ICT in Education

This chapter provides an overview of the economic literature on the effects of ICT in education and, in relation to this, the barriers to technological change that teachers (potentially) experience. For the literature on effects of ICT in education used in this chapter, only empirical analyses of ICT interventions in which the effect is computed on educational outcomes are included. These include cognitive tests, but also the motivation of students and non-cognitive outcomes such as computer skills. In this literature overview, only experimental research designs or quasi-experimental research designs are considered, as they are the only research methods that ensure that the conclusions can be causally interpreted. As mentioned, causal inference means that the researcher can be sure that the studied effect in the outcome measure (e.g. student performance) is solely due to the studied intervention (ICT in this case), and not to other factors indirectly influencing the outcome measure. The quasi-experimental designs that are included are: Regression-Discontinuity Design, the Difference-in-Differences Approach, the Instrumental Variables (IV) Approach and Statistical Matching (only when comparability between groups is shown), such that potential selectivity is controlled for (the importance of which is emphasized in De Witte and Rogge, 2014) (see the Appendix for an explanation of these research methods). Studies included in this review must have been written in English, and published in a peer-reviewed journal or an official working paper series.

Computers in education in general

In the economic literature, many of the previous evaluations of information and communications technology (ICT) are broad in scope, as they typically evaluate the impact of an ICT-enabling policy shift, e.g. increased ICT budgets for schools or households or the availability of free Wi-Fi on university campuses. These evaluations often study the general expansion of devices or connections. Here, the underlying assumption is that schools will use the additional budget for ICT efficiently. In these types of analyses, economists do not focus directly on evaluating educational technology, which is the focus of most educational studies, but more on general ICT increasing policies, that might not even directly influence educational performance. In many cases, the additional budget is positively related to an increase in ICT use, which is assumed to also lead to higher student performance. Empirical evaluations, however, have found very diverse results. Both positive and negative as well as zero effects are observed in the literature. With respect to the latter, Goolsbee and Guryan (2006) studied the impact of home internet connections on student performance and did not find a statistically significant result. In contrast, a large proportion of the studies on general ICT use in education find positive effects. For example, Machin et al. (2007) used an instrumental variable approach to identify the causal impact of ICT expenditure on the performance of students, as measured by the national test at Key Stage 2 (age 11), noting a statistically significant positive relationship. Punie et al. (2006) also found a positive relationship and conclude that there is evidence that ICT use improves the way of learning and individual student performance, although ICT use at home accounts for a major part of this relation. Lastly, Sosin et al. (2004) concluded that ICT has a small but positive effect on the performance of students. A negative relationship has, however, been observed by Leuven et al. (2007), who looked at the effect of computer subsidies on performance and concluded that the money is not used to buy computers, but rather to buy software and connect to the Internet, which seems to be used ineffectively. Furthermore, Angrist and Lavy (2002) found a negative relation between ICT and school performance. A potential reason for finding no or even negative effects could be that the ICT (budget) was not used effectively by schools. Another possible reason is that the ICT devices proved distracting to students (Beland & Murphy, 2016; Carter et al., 2016). However, since

the results are mixed, it is hard to draw any conclusions from these studies.

Lastly, there are two more recent studies in Peru for the One Laptop per Child Program, which found mixed results. Christia et al. (2012) studied the effect of the One Laptop per Child Program in rural Peru. This large-scale randomized program was implemented in 319 primary schools. Although computer access and use had increased substantially in the treatment group compared with the control group (who did not receive laptops), the authors found no evidence that there is an effect on motivation, enrolment and mathematics and language test scores. However, they did find small positive statistically significant effects for a general cognitive test, namely Raven's Progressive Matrices. Overall, the computers primarily appeared to be used for 'normal computer use' (word processing, internet browsing, games, music, etc.) and less frequently for didactic applications, which could potentially explain the lack of evidence on most outcomes. Beuermann et al. (2012) also studied the One Laptop per Child Program in Peru, but focused on primary schools in the Lima area. They also found that treatment children are better at using computers, but that there was no effect on objective and self-reported skills. A small positive effect on Raven's Progressive Matrices test was found, but only for children that did not previously have a home computer. In relation to the Peruvian study, Bando et al. (2016) studied the effect of replacing traditional textbooks with laptops (digital textbook provision) in high-poverty communities in Honduras in 2013 among 271 elementary schools with 9600 students. They found no effects of substituting textbooks with laptops, but argued that the policy might nonetheless be cost-effective, because textbooks are also very expensive and the cost of five books would be equivalent to investing in a laptop.

Computer directed vs. traditional classroom teaching

Another section of the literature on ICT focuses on the comparison of computer-directed versus traditional classroom teaching. A limited number of meta-analyses apply strict selection criteria with respect to the methodology used in the individual studies (Cheung & Slavin, 2012, 2013; Kulik & Kulik, 1991; Means et al., 2010) and show that, in general, computer-directed instruction does have small positive effects on student performance compared with traditional classroom teaching for both mathematics and language. However, these results may differ among students. MacGregor et

al. (1998), for example, showed that the effects on mathematics of a computer-augmented learning environment differ for students in regard to their preferences for how they learn, underlining the results of the studies that advocate a greater differentiation between students because of differences in learning styles. Tournaki (2004) concludes that, in general education, students benefit from drill and practices in the classroom, whereas students with learning disabilities do not benefit from this method. Becker (1990) studied the use of traditional classrooms compared with digital classrooms, but concluded that there are no statistically significant differences in student performance.

Other interesting interventions in education involving ICT and computers which are not covered in this chapter include, for example, Flipping the Classroom, MOOCs (Massive Online Open Courses) and quizzing (note that this is not an exhaustive list). However, these interventions are either primarily used in higher education (such as MOOCs) or there is no currently available causal evidence regarding the effectiveness of these interventions (such as for Flipping the Classroom). Therefore, these interventions are not part of this literature overview. Furthermore, it is plausible that no effects of ICT are found on student performance (e.g. Becker, 1990), but that the process has become more productive or efficient, leading to lower costs. However, to date, there have been no causal studies analyzing this aspect, and therefore these are also not part of this review.

Evaluating the contribution of digital learning tools

ICT proves particularly suited to providing individualized differentiation (hereafter, individualization), with its algorithms that allow for individual learning paths. Incorporating the differences in level, interests and learning styles between students is shown to improve students' motivation (Tomlinson, 2004), and neglecting these differences might lead to a decreased performance in certain students (Tomlinson & Kalbfleisch, 1998).

However, the effects for mathematics and language seem to be quite different, and differences in effects are also found in studies for developing ver-

sus developed countries, as Bulman and Fairlie (2015) demonstrate in their overview study. Overall, they find that the evidence of effects in education is mixed, at best. They conclude that, in general, studies with better methodological approaches (such as Randomized Control Trials (RCTs)) find small or null effects, where studies with methodological approaches of lower quality (such as multivariate and IV approaches) find larger and generally positive effects. They conclude that the only two areas where positive and statistically significant effects are found are studies in developing countries and studies that target mathematics rather than language.

However, in this overview study, significantly more studies are included than my inclusion criteria would allow, as they also include non-causal studies, where it is unclear how much of the found result is actually attributable to ICT. Furthermore, some more recent studies are missing from their overview. Therefore, *all* relevant experimental and quasi-experimental studies below will be separately briefly described in respect to mathematics, language and other cognitive and non-cognitive skills. In general, with regard to the educational sectors, there are slightly more studies on secondary than primary education, but studies on vocational education are underrepresented. The division of studies between western and non-western countries seems quite equal, so there is also representativeness from that aspect.

Experiments on mathematics skills

Evaluations of ICT-based individualization programs in mathematics range from general teaching to remedial programs and cover both general student audiences and students with learning disabilities. Overall, evaluation outcomes tend to be positive. It is worth noting that, in most studies, the control condition is traditional instruction without ICT, and this will only be mentioned if this is not the case. Burns et al. (2012), for example, show that computer-delivered mathematics fact intervention used three to five times a week results in greater gains for participating third- and fourth-grade students in the United States, compared with control students that did not or barely used this program. Furthermore, they show that significantly fewer of the students at risk for mathematics difficulties were still at risk after using a computer-delivered mathematics fact intervention. Similar results are found by Pilli and Aksu (2013) who studied the effects of computer-assisted instruction (CAI) for 55 fourth-grade students on mathematics per-

formance in Cyprus. They conclude that effects are found for multiplication and division, but not for fractions. Banerjee et al. (2007) reported the positive outcomes of an experiment with an ICT-based mathematics remedial program, introduced in primary public schools in two cities in India, which illustrates that the benefits of ICT-individualization are not confined to students from highly technologized societies. Arroyo et al. (2010) analyzed 250 seventh- and eighth-grade students in Western Massachusetts, the US, that used a digital skill drill method or traditional practicing on paper, 15 minutes per day following mathematics classes, for four days, and found a statistically significant positive effect of digital practicing. Barrow et al. (2009) also performed a randomized experiment in the US, among 1605 middle and high school students, and demonstrated that treated students perform significantly better on pre-algebra and algebra skills than their counterparts who received traditional instruction. Mo et al. (2013) studied the effect of the One Laptop per Child program in Beijing migrant schools in China, among 300 third-grade students, finding that there is a positive effect on computer skills and on mathematics scores. The program also increased the amount of time students use educational software in comparison with time spent watching TV. In another study, Mo et al. (2014) conducted a clustered randomized field experiment on the effects of an in-school computer-assisted learning program among third and fifth graders of 72 rural schools in Shaanxi, China. Students were offered the tool twice per week (40-minute sessions) throughout an entire school year and practiced their mathematics skills using computer-based games, which were tailored to their regular mathematics curriculum and were remedial in nature, providing practice at different levels. The authors concluded that the overall mathematics performance significantly increased by 0.17 standard deviations across all students. With respect to interpretation, an effect of 0.2 of a standard deviation is considered a small effect size, a standardized effect of 0.5 is considered a medium effect and 0.8 and over is considered a significantly large effect for interventions in education (Cohen, 1988). Christensen and Gerber (1990) studied the effectiveness of computerized drill and practice games on basic mathematics facts compared with an ordinal skill drill program (no game), and concluded that the game format is not beneficial for learning-disabled students. The oral and computer format seemed to work well, but did not show statistically significant differences.

Bartelet et al. (2016) studied the differential effect of an intelligent tutor-

ing homework system (ITS)[3] for mathematics skills among seventh-grade students in the Netherlands and found that the effects differ among achievement subgroups and mathematics domains. The control group in this case did not receive additional homework. Although the implementation of a non-compulsory supplementary tool leads to improvements in the overall mathematics skills of both low- and middle-achieving students, it is significantly less effective for middle achievers than for low achievers. This finding is especially true for the analyses divided among mathematics skills domains (numbers, proportions (fractions), and measurement). They also find that having access to the tool does, on average, have a statistically significant positive effect on students' proportions (fractions) skills, but not on their numbers and measurement skills. The difference in finding for fractions in comparison with Pilli and Aksu (2013) is most likely due to the age of the student in relation to the difficulty of the mathematics domain, as this is a more difficult domain than, for example, numbers. Haelermans and Ghysels (2017a) used the same experiment as Bartelet et al. (2016) to study the general mathematics effects of practicing with this homework ITS and concluded that the program is effective in increasing mathematics scores, and that this is primarily due to the adaptive individualized nature of the program. De Witte et al. (2015) analyzed the effect of a computer-assisted mathematics learning program, also in the Netherlands, and concluded that working with the program seems effective. The control group consisted of students that did not use the program or barely used it at all. However, they also found that schools with lower educational attainments are more likely to use such a program, which might lead to selectivity. Cabus et al. (2017) studied the effect of using an interactive whiteboard (SMARTboard) in class, and training teachers on how to use this among 199 seventh-grade students in a Dutch secondary school. The SMARTboard was used to differentiate learning materials among students, communicate more directly with students, provide students with additional learning materials when relevant and providing students with more independence in their learning process. They concluded that level differentiation in class, which was possi-

_____

3. ITSs are a specific type of computer-assisted instruction (CAI; a categorical name for programs that use technology to enhance cognitive achievement), which are thought to be particularly effective because they are adaptive and interactive, and are often used in remedial education contexts.

ble because of the efficient use of the SMARTboard, significantly increased mathematics proficiency by 0.1 of a standard deviation, which is a small effect. Lastly, Falck et al. (2015) studied the effects of different computer use in the international TIMMS (Trends in International Mathematics and Science Study) data, exploiting within-student and between-subject variation, and found that the zero effects that are often found are a combination of positive effects of using computers to look up information, and negative effects of using computers to practice skills, although the effect of the latter appears to also be related to the teacher, underlining the critical importance of the teacher (as will be more extensively discussed later in this chapter).

Another study from developmental psychology used a randomized controlled trial to enhance numerical cognition of primary school children with dyscalculia with a computer-based training program (compared with delayed access to this program) and also found positive effects of the program (Käser et al., 2013).

Experiments on language skills

In the literature on language tools, we see only a limited number of experimental studies and very few of them have found positive effects. Borman et al. (2008), for example, conducted a randomized field trial in the United States to study the effects of computer-based training (FastForWord) on language skills, and concluded that this program did not help students improve their language skills. The control condition received other instructions and activities that were not related to literacy instruction. Similar conclusions are drawn by Rouse and Krueger (2004) and by Given et al. (2008), who both studied the effectiveness of the same program on student performance in the US. Potocki et al. (2013) also found no statistically significant results when using a randomized experiment to study the effects of computer-assisted training for language comprehension (compared with computerized instruction for decoding skills) in France. Interestingly, the only source that concludes that small positive effects are likely is the meta-analysis by Cheung and Slavin (2012). This finding cannot be explained directly, as the article does not show what the effect sizes are per study, nor how the authors estimate the overall small but positive effect. Furthermore, an earlier meta-analysis by Slavin et al. (2011) showed no statistically significant results. A very recent study by Ghysels and Haelermans (2016) also found

small, but positive, effects of practicing with an online homework tool for seventh-grade students on spelling performance in the Netherlands. Here, the control condition was computerized instruction with other aspects of language learning. These effects are primarily due to low-performing students. The authors attributed the positive effect they found, in contrast to all other studies that found no effects, to the fact that spelling is an easier to automatize skill, compared with, for example, reading.

Experiments on both mathematics and language skills

Some studies look at both mathematics and language simultaneously. Carrillo et al. (2010), for example, used a randomized experiment to study the effect of computer-aided instruction for mathematics and language in Ecuadorian primary schools (the control schools only received the computer-aided instruction later, following the experiment). They found a positive and statistically significant effect on mathematics performance, but a negative non-significant effect on language performance. The effects appear to be much higher for well-performing students. The authors believe the positive effect for mathematics is due to offering a good combination of hardware, software (APCI platform (Aprendizaje Personalizado Complementario e Interconectado – the Personalised Complementary and Interconnected Learning software tool)) and teacher training. However, they do not have an explanation for the lack of findings for language.

Experiments on other cognitive and non-cognitive skills

Besides the studies on mathematics and language described above, there are numerous studies that look at the effects on other cognitive skills, as well as on some non-cognitive skills. Linden (2008) studied the effect of a computer-assisted learning program on student achievement among third- and fourth-grade students in India. If the program was used as a teacher substitute, student achievement in the treatment group was significantly lower than before, implying that students learn less from the program than from the teacher. If the program was, in contrast, used as a complement to the teacher, a positive but non-significant effect was found. This study shows the importance of thinking about how new technologies should be implemented in schools, and how they interact with existing resources. Bar-

rera-Osorio and Linden (2009) studied a randomized experiment among 97 schools and more than 500 students in Colombia where computers from the private sector were donated to public schools for teaching language. They found no effect of this donation on test scores and other outcomes, which they attribute to the fact that the computers were not effectively incorporated into the educational process, despite the fact that teacher training was provided. Haelermans et al. (2015b) studied the effect of digital differentiation, using a randomized experiment among 114 eighth-grade students in the Netherlands. Digital learning materials were made available to (all) the students in class, and half were asked to study at their own, adapted level, whereas the remainder studied at the mean level. The results show a statistically significant effect of digital differentiation on the students' biology grades.

Another study from cognitive neuroscientists on the effects of working memory training for children in primary school in Sweden also used computerized training (Klingberg et al., 2005). This randomized controlled trial used a small sample size (around 50 children), but was highly intensive, and found positive effects of the program on the working memory of the children. The control group also received computerized training.

Other studies have used a quasi-experimental design and, in most cases, rely on existing databases to study the effects of ICT in education. Christia et al. (2014) followed more than 7000 secondary students between 2001 and 2006 in the upper-middle income country Peru. Using a difference-in-difference approach, they found that no evidence that increased computer and internet access has a statistically significant effect on repetition, dropout and initial enrolment. Several robustness checks confirm these findings. The authors do not offer any explanations as to why they do not find any statistically significant results, but this could be another example of the use or introduction of ICT being studied, with no account taken of how the computers are actually used in the school. Fiorini (2010) uses quasi-experimental methods to study the effect of home computer access and use on primary school-aged children's cognitive and non-cognitive skills, based on an Australian dataset (LSAC, Longitudinal Study of Australian Children). The author found that there is a positive effect of computer time on cognitive outcomes, but that the effect on non-cognitive outcomes is mixed, depending on age and score.

Cost-effectiveness of digital learning tools

Unfortunately, only one or two of the abovementioned studies on the effects of digital learning tools also look into the cost-effectiveness of these tools. In many of the studied situations, ICT was provided by the government or by direct subsidies. However, if schools consider investing in ICT themselves, it is not only important to know the potential effects, but also the subsequent cost, given the size of the effect. In other words, schools would want to know if it is worth investing in this digital learning tool. The few studies that do present a cost-effectiveness analysis can roughly be divided into western and non-western countries. Linden (2008) and Mo et al. (2013) both estimate the cost of their student intervention to be about 5 USD per tenth of a standard deviation increase in student performance, in India and China, respectively, and Banerjee et al. (2007) estimated the costs of their CAL intervention in India to be around 15 USD per student per year. Bando et al. (2016) also performed a cost-effectiveness analysis of a book-replacing laptop program in Honduras, and concluded that the program would be cost-effective if only five books were replaced by their digital counterparts, despite the fact that they did not find an effect. They base this conclusion on estimations of the gains of replacing textbooks, the digital literacy premium, and the potential wage premium of 1% (a conservative number) for the increase in digital literacy.

Two studies that took place in the US also estimated cost-effectiveness, and found much higher costs than the studies in non-western countries. Rouse and Krueger (2004) estimated the costs to be 770 USD per student per year, if one adult (teacher) could handle 40 students at the same time. However, they did not incorporate the indirect costs of having to hire that adult and needing to spend a large amount of time on this program into this figure. Barrow et al. (2009) compared the costs of using CAI with reducing the class size to 14 students, and concluded that the costs were approximately the same (218 USD per student per year for the CAI if there were 30 students in the class in the lab versus 198 USD per student for a class size reduction from 23 to 14 students). Both these US studies incorporated the costs of setting up the computer lab, including the devices. A study in the Netherlands by Haelermans and Ghysels (2017a) concluded that the cost of an adaptive digital computer tool for students would be around 25 euro per student per year, although they did not consider the costs of the computers as their study looked into a homework tool and all students had computers

or could use the ones already present at school. This makes the latter less comparable to the cost-effectiveness analysis of the other western studies.

In general, the cost-effectiveness analyses executed in the above described studies generate very similar results, once studies conducted in western countries and non-western countries are compared as separate groups. Unfortunately, compared to all studies that analyze the effectiveness of ICT and/or digital learning tools in education, only few include a cost-effectiveness analysis. Therefore, the most important conclusion regarding this topic is that more research is needed to draw reliable conclusions on the cost-effectiveness of ICT in education.

Teachers' barriers to (technological) change

Despite the fact that technology in education has been around for quite some time, and that specific digital learning tools have shown to be effective for student learning in mathematics, but not for language training (except in non-western countries), technology adaptation in education has not, to date, evolved noticeably. Although most students now have access to computers, both at school and at home, the OECD also argues that students' performance has not increased between 2000 and 2012. The OECD argues that the performance of students in countries where ICT is used in the classroom is mixed and that the positive effects are not obvious (OECD, 2015). This raises one important question: if technology is present in classrooms, but the OECD does not observe an overall increase in educational performance, whereas studies on specific digital learning tools do show positive effects, at least for mathematics, then why is the educational sector (including teachers) not more interested in learning more about why technology in education is not being used optimally, and what kind of ICT tools in education contribute to student performance the most?

The answers to this question can partly be found in the body of literature on changes in education. Although virtually none of the studies provide causal evidence, the relatively large body of literature on changes in education is reasonably informative concerning the barriers to change in the

educational sector, and in particular to the introduction of technology in education. This literature is by no means new, but apparently the situation in the educational sector remains much the same.

In the literature on (resistance to) changes in education, and specifically in the literature where barriers to changes are discussed, one can roughly distinguish two types of teachers: 1) those that do not want to change, and 2) those that are willing to change but do not know how or feel insecure about the change. Ertmer (1999) has defined these as two types of barriers to change, namely first- and second-order. First-order barriers are extrinsic to teachers, such as equipment, time, training, support, etc., whereas second-order barriers are intrinsic and are more related to teachers' beliefs regarding technological change. First-order barriers are more related to the abovementioned second type of teacher, whereas the first type of teacher has more second-order barriers. It is important to note that, to date, there does not seem to be a clear relation between other observable teacher characteristics and the susceptibility, willingness and confidence of the teacher in using ICT (effectively) in class.

In regard to teachers that do not desire change, with second-order barriers, Beeby (1966), writing decades ago, stated that the ability of teachers to promote change is the most important factor in educational change in general. Teachers that do not desire change might experience a lack of clear goals, a lack of understanding and acceptance, might be the product of a system that is not very willing to innovate, or might be overly isolated in their own classroom, which slows innovation down (Beeby, 1966). Fifty years on, it seems as if nothing has changed. The arguments regarding why a teacher would be resistant to change still apply as of 2016. Diamond Hicks (2011) also mentions the same types of resistance to change by teachers, and Freilich Hjelle (2001) concludes that teachers are expected to blindly accept changes to their everyday practice, whereas their expertise on the topic has never been requested or taken into account.

As regards the teacher that wants to change but does not know how, the literature shows that teachers need consistent support and extensive training in order to believe that they are able to properly use the educational technology (Demetriadis et al., 2003). Pelgrum (2001), in his article on obstacles to ICT integration in the classroom, shows, based on a worldwide assessment, that, according to the educational practitioners, besides the availability of computers (which obviously could have been an issue in

the nineties), the teachers' lack of knowledge and skills and the difficulty to integrate the technology into instruction are the two major obstacles. In her review of the literature on factors that influence teachers to use ICT, Mumtaz (2000) lists the following issues that prevent teachers from using technology, namely a lack of experience, support, specialists, computer availability, the time to successfully integrate the technology in the curriculum, and financial resources. These factors are also noted by Keengwe et al. (2008), Cafolla and Knee (1995) and Hardy (1999). Bingimlas (2009) more or less draws the same conclusion from his review study, in which he shows that a lack of confidence, competence and access to resources constituted the main barriers for teachers that had a desire to use technology in education. Rakes and Casey (2002) emphasize that the personal concerns of teachers should explicitly be addressed when implementing technological change. These concerns, which might differ depending on the stage of concern a teacher is experiencing (this concern could be awareness, which is an early stage, but could also be not knowing how to manage it all, which is a later stage), can be addressed by appropriate training and allocating sufficient time to the change.

Change management literature

All of the abovementioned aspects come neatly together in the literature on change management. Although this strand of literature generally focuses on the private sector, which lies outside the scope of this report, the conclusions are also very relevant for the public sector (one literature review by Fernandez and Rainey (2006) specifically focuses on the public sector), such as education, and it is here worth briefly mentioning some aspects. First of all, the literature on change management shows that it is important to acknowledge that the failure of an ICT project is, in most cases, not due to technical reasons, but rather to non-technical ones, such as the non-acceptance of the solution, skills problems, communication problems, etc. According to Kotter (1995), the implementation of changes often fails due to one of eight factors: not establishing a sufficient sense of urgency; not creating a sufficiently powerful guiding coalition; lacking a vision; under communication of the vision, not removing obstacles (e.g. opposed individuals); not creating and celebrating short-term wins; contrastingly, declaring victory too soon; and, finally, not anchoring changes in the organization's culture.

In their literature review on successful organizational change in the public sector, Fernandez and Rainey (2006) have created a similar list of critical factors for successful organizational change: ensure the need, provide a plan, build internal support for change and overcome resistance, ensure top-management support, build external support, provide resources, institutionalize change, and pursue comprehensive change. These aspects are, in many respects, similar in nature to the barriers to technology adaptation in education seen above.

Conclusion

From the abovementioned literature on effects of ICT in education, we can draw one or two conclusions. First, general investments in ICT in education without a specific investment aim of how to use ICT in education (effectively) provide mixed results, at best. However, studies on the effects of computer-assisted instruction versus traditional classroom learning do find positive effects, albeit very small, and cost-effectiveness is not analyzed in these studies. This means that the way in which the computer-assisted instruction is developed and implemented is also very relevant and not just the introduction of ICT in itself. Furthermore, studies on specific digital learning tools, primarily focusing on improving mathematics and/or language performance, also provide mixed results. In general, these studies find positive effects for interventions taking place in developing countries, and interventions involving mathematics in all countries, but not for interventions focused on language training. Although, in most cases, no direct explanation is given in studies that find a specific positive or negative result, it seems that most studies on language look at reading and text comprehension. This leads researchers to believe that perhaps the positive effects in mathematics are due to the fact that mathematics (addition, multiplication, etc.) consists of more easily to automatize skills than the aspects of language that were studied.

However, the question could be asked as to why ICT in education is not, in general, used that extensively, or at least not very efficiently, while CAI shows positive effects over traditional classroom learning and adaptive digital learning tools are proven to be effective for mathematics. The literature

shows that there are many barriers to technological change for teachers, which might explain why technology adoption in schools has not, to date, lived up to expectations. The literature shows that teachers are either resisting the technological change in general, due to, for instance, their internal beliefs, or do not know how to apply the technology effectively in class, due to factors such as a lack of time, knowledge or training. The closely related literature strand on change management offers some useful suggestions on how to implement changes in public sector organizations such as schools.

Chapter 3.
The Research Context

This chapter describes the research context in which the experiments and quasi-experiments discussed in the next chapter take place. First, the main characteristics of the Dutch educational system that are relevant to understand the setting of the experiments in the next chapter are explained. Following this, Dutch national education policy with respect to ICT is briefly discussed, followed by some statistics regarding the use of ICT in education in the Netherlands. Lastly, a comparison is made between the Dutch and Swedish educational systems, in order to draw policy recommendations from the Dutch experiments for Sweden in the concluding chapter of this report.

Dutch educational system

In the Netherlands, full-time school attendance is compulsory for children between the ages of five and 16. However, the majority of students have already begun attending education by the age of four. From the ages of 16 to 18, students need to attend school at least part-time and sometimes full-time, depending on their type of school.[4]

The Dutch educational system has some key characteristics that are different from many other countries. The first peculiarity is the tracking system of students in secondary education. In this respect, the Netherlands resembles some other European countries such as Germany and Switzerland. In

---

4. Note that this is different from Sweden, where in theory, upper secondary education is voluntary for children to attend from age 16 onwards (although almost all children do attend some form of education between 16 and 18).

the Netherlands, pupils attend primary education between the ages of four and 12 and secondary education until a higher secondary degree is obtained. Within the Dutch educational system, one can distinguish five different levels of secondary education: practical training, prevocational secondary, vocational, general upper-secondary and pre-university. Pupils enter a level of education in the seventh grade, which is the first year of secondary education in the Netherlands, based on a standardized national test and a recommendation given in elementary school (i.e., ability tracking). Depending on the level of education, secondary education takes four, five or six years to complete. Practical and prevocational training take four years to complete, with vocational at least another two years in addition to prevocational, with a maximum of four years in addition to prevocational. General upper secondary education takes five years, and pre-university education takes six. Within tracks, students have to choose specialized courses. In upper secondary school, they can decide between a culture and a nature track in general upper secondary and pre-university education, and between a health, economics, agriculture and technical track in (pre-)vocational education.

A second characteristic of the Dutch system is grade repetition. Because of the tracking system, students have to be able to meet the level of the track for *all* courses they follow. If they perform badly at one or two classes, they are not allowed to continue to the next year, leading to grade repetition, being placed back a level to a lower track or, on some occasions, even dropping out. However, students hardly move from a lower track to a higher track, once they started in a certain track. On the other hand, grade repetition (often combined with stepping back one track) is very common. Almost half of the students repeat a grade in primary or secondary school (Van Vuuren & Van der Wiel, 2015), of which almost two thirds originate from secondary education. Grade repetition may often lead to unlawful absence, which, in turn, may on some occasions lead to school dropout.[5]

A third characteristic of the Dutch educational system is the concept of "free school choice" and "freedom of education". The former element of the Dutch education system is comparable to US charter schools, or Swedish free schools, that can be attended by choice; this means that students do not have to live in a particular catchment area in order to be eligible to at-

---

5. The dropout rate in the Netherlands was 8% in 2016.

tend (Imberman, 2011). In contrast, for primary school, almost all students attend the nearest school. For secondary school, students receive secondary school-level advice from their primary school. In general, the same educational track is offered by several schools, so, together with their parents, students can freely choose a particular secondary school. However, in practice, they often choose a school nearby, and in many cases this is the nearest school, but this is not always the case. This also depends on the tracks that are offered at the nearest school and which school their friends attend. As for "freedom of education", this implies that public and free schools are statutorily equal and are funded equally by the government (see below) so parents are free to choose between public schools and free schools. Free schools do have more freedom in how they organize their education, thus, in that sense, they could be considered as private schools. However, 'real' private schools, which are not funded nor regulated by the government, barely exist. Only a handful of those schools exist among a total number of more than 600 secondary schools. Free schools can have different denominations, ranging from denominations of a religious nature such as Roman Catholicism or Protestantism, a reform pedagogical nature such as Montessori or Waldorff schools, or with a cultural nature, such as Chinese schools. Freedom of education also means that everyone with a viable plan has the right to start a new school, if can be argued that there is need for this new type of school, i.e. enough children are willing to attend. There is no indication to believe that ICT use might be different in these free schools in comparison to public schools.

The funding system of (secondary) schools in the Netherlands is that schools receive a lump-sum payment from the government every year. The amount that schools receive is basically dependent on the number of students, with a correction, in such a way that schools that have a high proportion of students with a low socio-economic status receive a higher budget. Within the existing legal framework, the allocation of this budget among the several resources is the decision of the school. It is worth noting that the lump-sum budget excludes (large and discontinuous) payments for housing infrastructure (the latter are provided by the municipal and central government). In most cases, the municipality is responsible for providing the accommodation of the school, although the school itself is responsible for the maintenance of the building. In practice, this means that the school is granted a certain amount of finance from the municipality to fulfil this re-

sponsibility. Building new accommodation often only takes place when the number of students has risen substantially in the previous years and is solely based on the number of students. Aside from the lump sum, there are no other general sources of income to the schools. Schools are allowed to ask for a voluntary parental contribution, but this contribution can only be used for the benefit of the student (e.g. for school trips).

At the end of secondary education, students have to pass national exams for all their courses in the final year of the curriculum. These national exams consist of written school exams, oral school exams and national standardized comprehensive written exams. The final grade for a subject is the average of the school and comprehensive exams. The average graduation grade of a student is her average grade over all subjects. In the grading system in the Netherlands, grades range from 1 to 10, where 1 is the lowest possible grade and 10 the highest. In practice, however, grades almost always vary between 4 and 9. Grades also usually have one decimal place, which is rounded off to the nearest whole or half number. Grades of 5.5 and upwards are considered to be the lowest pass grade. In the exams, students are allowed to obtain one 4 or two 5's, as long as these are compensated by higher pass grades in such a way that the average grade over all exams is at least a pass of 5.5 out of 10.

Passing the final exams and obtaining a diploma is a necessary requirement for enrolment in an institute for higher or vocational education. In the Netherlands, it is possible for a student to fail one of the exams. Students are allowed re-sit exams for this course. If students fail these re-sits or if they fail more than two exams, they cannot graduate that year. Such students will then have to repeat the graduation year. Often, these students will sit the national exams for a lower educational level the next year.

School quality is assessed by the national Education Inspectorate, which assesses each school at least once every three years based on performance in school and national exams, along with school climate. Schools that perform below an acceptable level will be assessed annually and need an improvement plan that is approved by the Education Inspectorate.

In the Netherlands, schools are free to choose their methods and formats of teaching. However, many schools gravitate towards methods that allow the children to have some independence in the process of teaching and learning. In order to ensure comparable quality in all schools, the government sets minimum standards for the level that all students need to have achieved by the time they leave primary school and by the time they leave

(a track of) secondary school. In this sense, schools are also free to choose to invest in ICT if they feel that this best suits their methods and forms of teaching.

Dutch educational ICT policy

Before, there was a clearer Dutch national educational ICT policy than in recent years. Fifteen years ago, at the beginning of the 2000s, there was a national policy on ICT in education to increase ICT use. This policy was seen as successful, given the widespread adoption of computers in education. Back then, there was already the challenge that teachers might not know how to use the ICT and should be given training, as well as the potential lack of time and the lack of knowledge on how to implement the ICT in the didactical process (Inspectie van het Onderwijs, 2001). The advice was to keep investing in ICT in education. However, another report by the Education Council in 2006 emphasized that if schools want to invest in ICT, they have to use their lump-sum amount for this, and therefore they should be very well informed about the effectiveness of certain types of ICT (Onderwijsraad, 2006). They also stated that the effectiveness of investments in ICT was not undisputed, and called for more controlled (and financed) experiments on ICT in education to study the effectiveness. In so doing, schools could participate in these experiments and have the way they use it studied, and through this, be better informed on the effectiveness of ICT. Although schools had to finance their own ICT investments, the government did invest in public organizations for education and ICT, such as NRO (the Netherlands Initiative for Educational Research) and "Kennisnet", which literally means 'web of knowledge'. Kennisnet aims to provide for a national ICT infrastructure in education, advise sector organizations, and share their knowledge with educational institutions in primary, secondary and vocational education. They also publish annual statistics on the use of ICT in education (see next section). NRO is part of the Dutch Scientific Organization and is responsible for distributing the funds for research on education, among which research on ICT in education, and for utilizing the knowledge in practice, by making sure research results reach those that should use it to base their decisions upon.

Currently, there does not seem to be a national policy regarding directly stimulating ICT in education, although the government is still in favor of greater ICT use. When national budgets needed to be cut, starting halfway between 2000 and 2010, before and during the financial crisis, national policies on ICT in education gradually declined. There are currently no specific national subsidies to stimulate ICT use in education, and even the subsidies for the public organizations such as Kennisnet were drastically decreased in 2013 (Ministerie van Onderwijs Cultuur en Wetenschappen, 2013). However, it is, to some extent, part of some governmental policies, such as national policies on teachers and on our future education (Education 2032), both of which focus on preparing the educational sector for the future. Of course, ICT is mentioned in these policies, as ICT and ICT skills are indispensable for the future, but more as a means to an end and not as a goal in itself.[6]

In line with freedom of education and the way the funding system works, schools are free to choose the method they want to use in their teaching, and to spend more or less money on this aspect. Each school individually decides which ICT tools to use, if any, and in what way (at school/at home). This predominately results in varying ICT usage in the three educational sectors considered here, as will be discussed in the next section, but also in very different ICT usage between schools and even between teachers (as teachers are often also relatively free to apply whatever they want to use in their classroom).

Moreover, (effective) ICT use is an issue in many schools, and many schools work with digital learning materials, whiteboards, and so on (see next section). The Netherlands even has initiatives such as iPad schools, where there are no traditional learning materials, just the iPad. However, these schools have come in for widespread criticism, as the general consensus is that digital devices are more a means to an end and should not be the goal in themselves.

Overall, one can conclude that educational ICT policy currently only takes place at a school level, or perhaps at the level of the governing body of the school (where multiple schools frequently belong to the same governing body) and not really at the national level.

---

6. www.onsonderwijs2032.nl.

ICT use in Dutch education

As discussed above, the public organization for education and ICT, *Kennisnet,* is the frontrunner organization when it comes to providing annual statistics on the use of ICT in education in the Netherlands. Each year, they publish a report named '*Vier in balans monitor*', which literally means 'four in balance monitor'. Four stands for the four inseparable elements that need to be balanced in order to make ICT in education successful: vision, expertise, content and application, and infrastructure.

The latest '*Vier in balans monitor*', published in 2015 (Kennisnet, 2015), contains statistics from the school year 2014/2015, and draws comparisons with the two preceding school years. The 2015 publication is based on questionnaire answers from more than 400 primary school teachers, almost 400 secondary school teachers, almost 250 vocational teachers, more than 200 primary school managers, almost 100 secondary school managers, and almost 80 vocational school managers, totalling more than 1000 teachers and around 400 managers. The monitor has an appendix which describes the origin of the respondents per sector of education, with respect to number of schools, type of contact, average age and gender. The sample does not seem to be selected randomly, and the authors do not perform an analysis to check the representativeness of the sample. However, a comparison with the national averages at least shows that the numbers on age and division of gender do represent the national averages for that sector.

The figures presented below are direct copies from this '*Vier in balans*' report, although captures have been translated into English.

Figure 3.1 presents the average number of students per computer at school between 2007–2008 and 2014–2015. Here, we can see that the number of students per computer has decreased from approximately six in 2007-2008 to approximately four in 2014–2015. This number is slightly lower for vocational education, and slightly higher for primary education. The dots in this figure are the actual number for each year, while the line is the linear trend line through these dots.

Figure 3.2 shows the number of hours per week that ICT is used in class, as decided by the teacher, for the three sectors for the school years 2012–2013 and 2014–2015. It shows that the share of classes where ICT is being used in class for less than five hours per week has decreased for all three sectors, whereas the share of classes where ICT is used more than 15 hours per week

Figure 3.1   Number of students per computer.
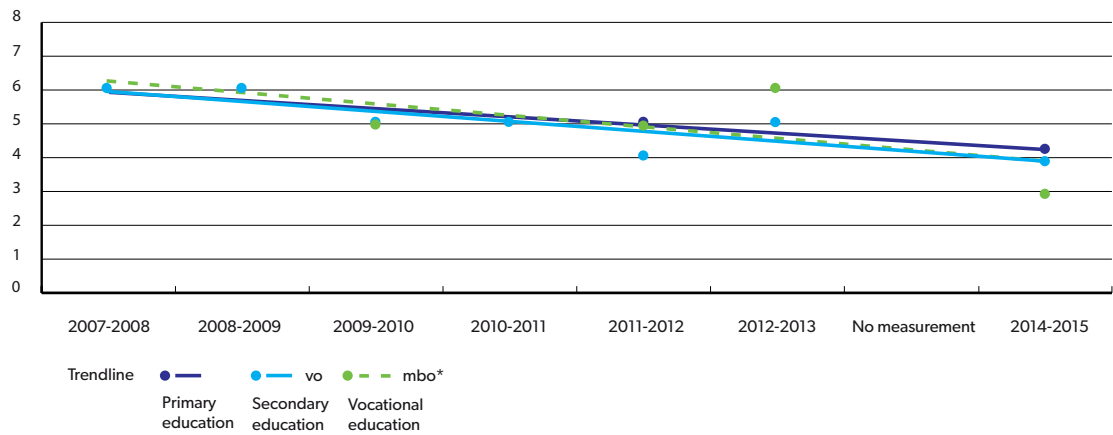
Source: Vier in balans monitor 2015, page 64 (Kennisnet, 2015), captions translated into English.

Figure 3.2   Number of hours per week ICT is used in class.
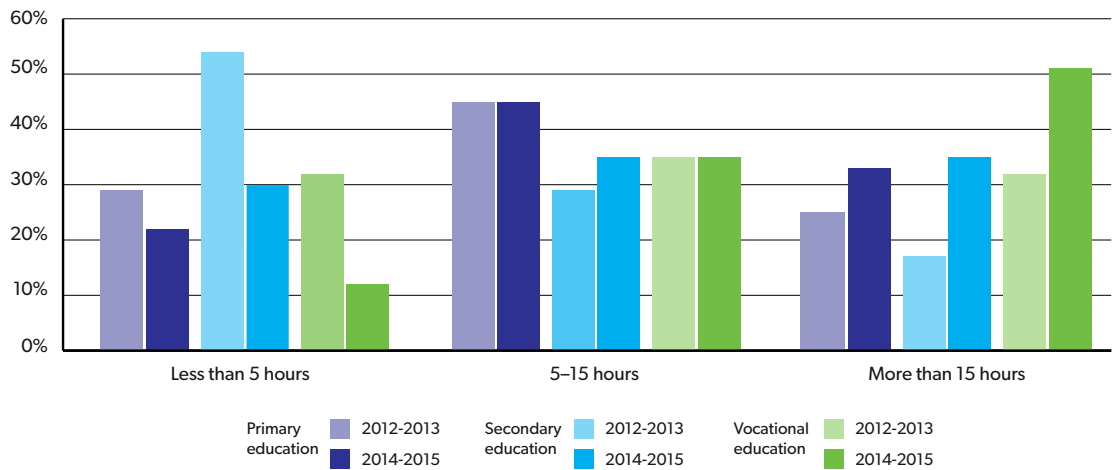
Source: Vier in balans monitor 2015, page 34 (Kennisnet, 2015), captions translated into English.

has increased. Secondary education shows the largest differences in decrease and increase. The share of classes where ICT is used between 5 and 15 hours is almost equal in the two school years that are being considered. In general, it seems that the time ICT is used in class has shifted from between zero and 15 hours to between 5 and 15+ hours.

Figure 3.3 shows the digital learning materials that are used in primary education, arranged by highest use, i.e. at least weekly. Figure 3.3 shows that 80% of the teachers use methods-related software and interactive practice material at least weekly in primary education, whereas videos/movies and printed text files are used at least weekly by more than 70%. What is striking is that e-books, digital tests, and interactive websites are rarely used on a regular basis.
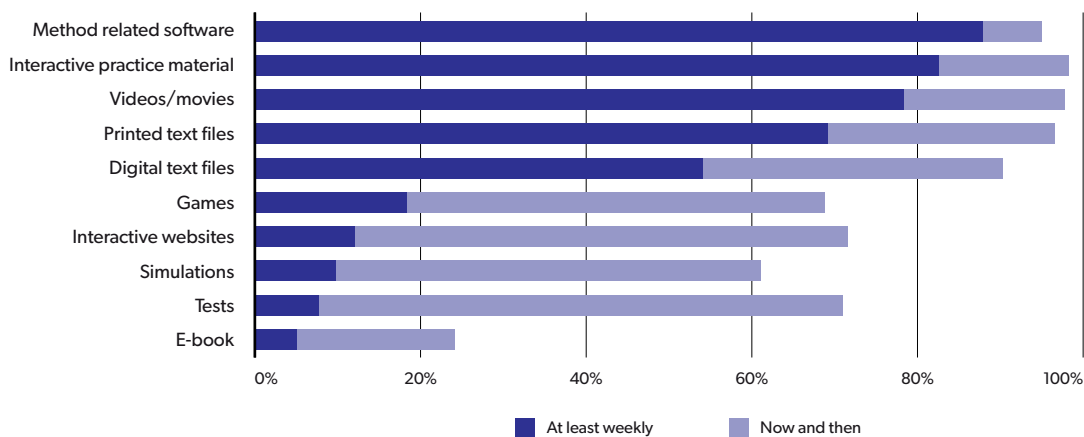
Figure 3.4 shows the digital learning materials that are used in secondary education, arranged by highest use, i.e. at least weekly. Figure 3.4 shows that more than 50% of the teachers use printed text files, digital text files, and video/movies at least weekly. It is striking that method-related software and interactive practice material are only used more than once per week by around 40% of the teachers, whereas these had the highest usage in primary education. Similar to primary education, tests, simulations, e-books and interactive websites are rarely used on a regular basis.

Finally, Figure 3.5 shows the digital learning materials used in vocational education, arranged by highest use, i.e. at least weekly. Figure 3.5 shows a very similar pattern to that of secondary education in Figure 3.4. Digital text files, printed text files and video/movies are used regularly by more than half of the teachers, followed by method-related software and interactive practice materials.

Figure 3.6 presents the origins of the digital material that is used in class. Most of the digital learning material that teachers use is provided by (the publisher of) the learning method, followed by Google, for both primary and secondary education. Primary education teachers also use various digital materials from digital video sources and educational websites, whereas secondary teachers develop the material themselves more frequently, or receive it from colleagues. Very few teachers take the material from social media or the national wiki learning materials website.

Figure 3.7 shows the classroom uses of digital applications. It is striking that using ICT to provide student instruction and using ICT for practice only come fourth and fifth in this list. It seems that ICT is generally used

Figure 3.3   Digital learning materials used in primary education.

Source: Vier in balans monitor 2015, page 58 (Kennisnet, 2015), captions translated into English.

Figure 3.4   Digital learning materials used in secondary education.
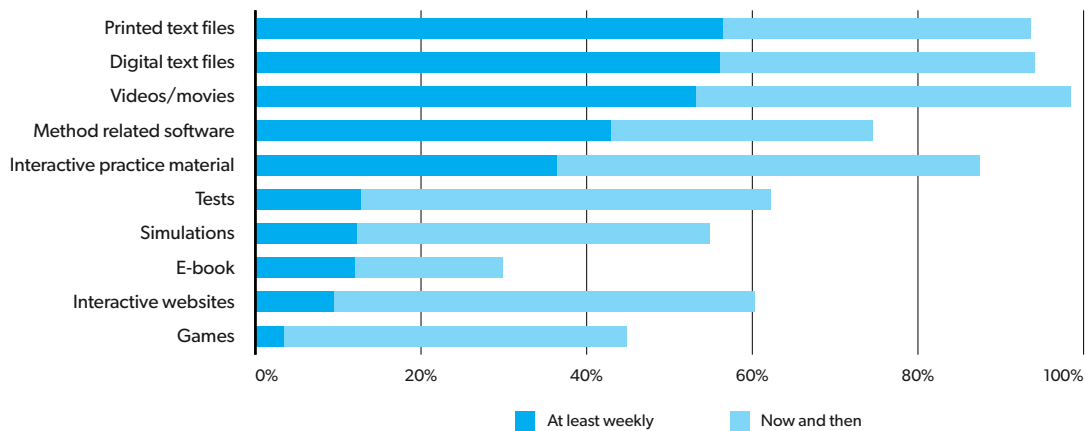
Source: Vier in balans monitor 2015, page 58 (Kennisnet, 2015), captions translated into English.

Figure 3.5   Digital learning materials used in vocational education.
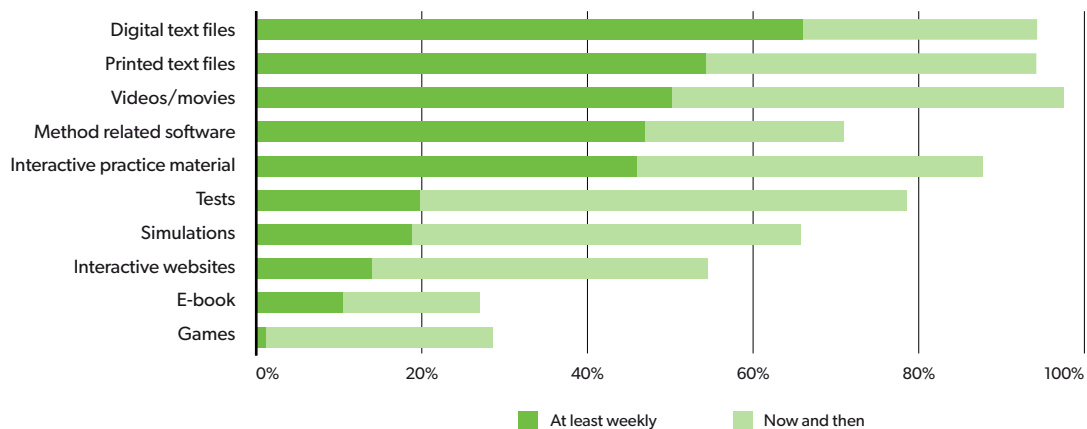
Source: Vier in balans monitor 2015, page 59 (Kennisnet, 2015), captions translated into English.

Figure 3.6   Origins of the digital material.

Source: Vier in balans monitor 2015, page 60 (Kennisnet, 2015), captions translated into English.

Figure 3.7   Used digital applications.

Source: Vier in balans monitor 2015, page 35 (Kennisnet, 2015), captions translated into English.

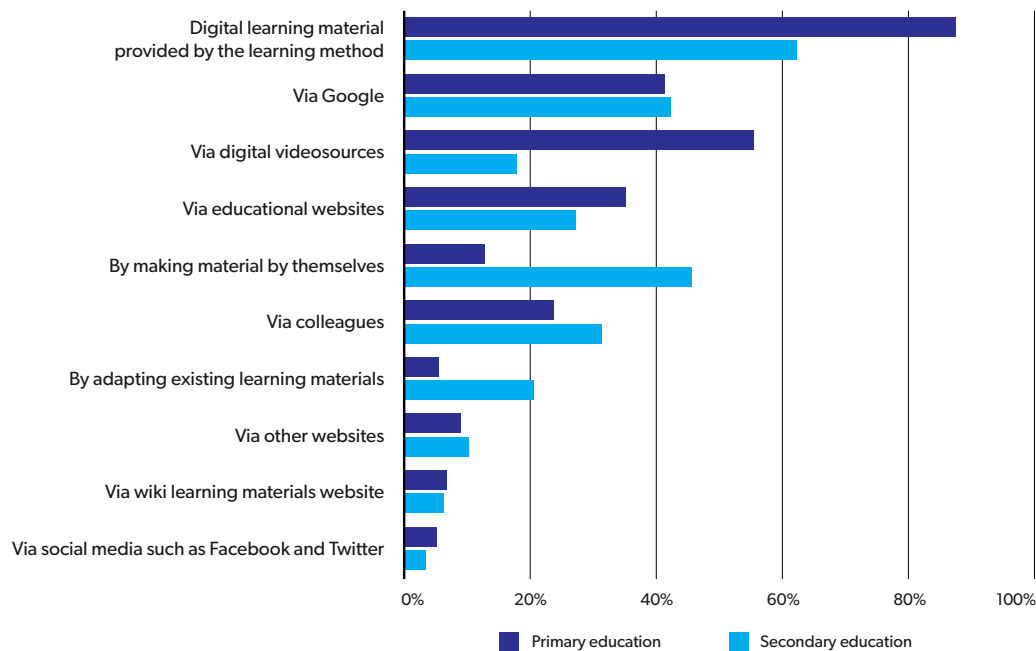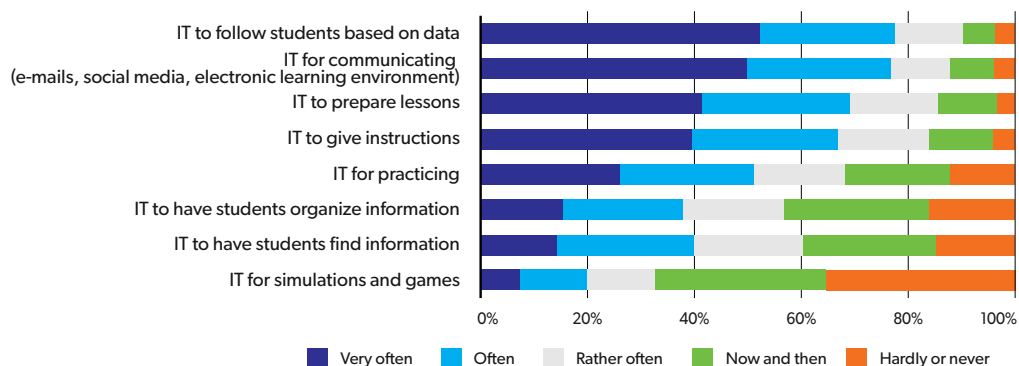for administrative purposes. Figure 3.7 shows that teachers typically use ICT to follow students, to communicate, and to prepare lessons. Figure 3.7 also shows that they rarely use ICT in class for simulation and games, to have students find information, and to have students organize information.

Finally, Figure 3.8 shows which digital information systems are used most often in the three sectors. Primary schools principally use the student performance system and the digital testing system, followed by the student registration system. Secondary and vocational education, on the other hand, generally uses the student performance system, the scheduling system, the absenteeism system, the student registration system and the electronic learning environment. This shows the very different uses of digital information systems between primary education and the other two sectors. The curriculum development system is only moderately used by vocational education and is barely used at all by primary and secondary education.

Overall, ICT is used in many schools in many different ways, but it seems that ICT is used more frequently for administrative reasons than for actual teaching purposes.

Figure 3.8   Use of data from digital information systems.
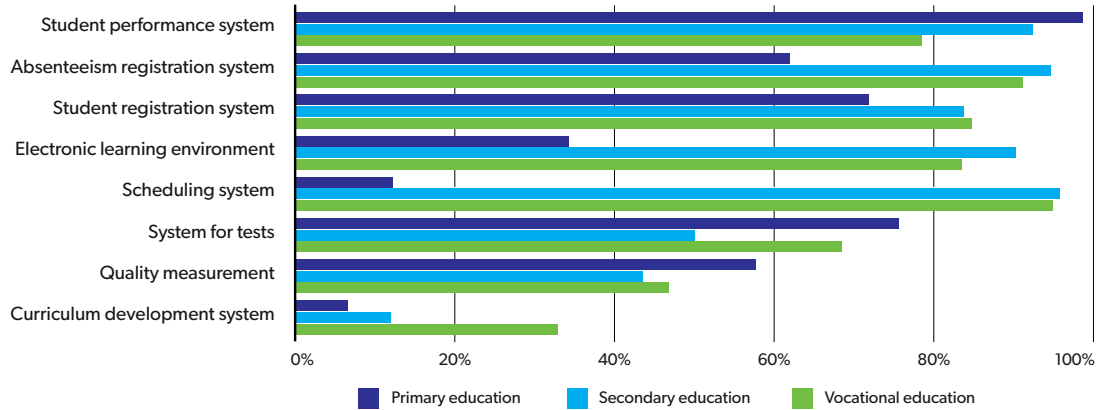
Source: Vier in balans monitor 2015, page 72 (Kennisnet, 2015), captions translated into English.

Comparison between the Netherlands and Sweden

In the characteristics of the educational system, the Netherlands and Sweden are not actually that different. The key characteristics of the Dutch system are also present in the Swedish system, namely the freedom of school choice, the tracking in (upper) secondary school in academic and vocational tracks, the public funding of free schools, and the freedom of schools to decide which learning materials and types of learning in which to invest. Another similarity is that of the status of teachers, as, in both countries, teaching is not seen as an attractive career, teacher appraisal is underdeveloped, and teachers have to perform all the administrative duties themselves.

However, there are also some differences between the systems, for example with respect to the compulsory nature of upper secondary education (not compulsory in Sweden) and the lump-sum budget being financed by the national government in the Netherlands, but in Sweden by the local governments, where the amount differs per local government. Another difference is the occurrence of grade repetition, which is highly uncommon in

Sweden, yet very common in the Netherlands. However, it should be noted that in the scientific literature, there is little evidence that repeating a grade is effective in rectifying the learning backlog that led to the repetition in the first place (Goos et al., 2013).

A similarity with respect to the national ICT policy is that, currently, in both countries no national ICT policy seems to exist. Given that, in both countries, schools receive a lump-sum payment, it seems to be an issue of where the responsibility is placed much more at the school than at the national level.

The Netherlands and Sweden are largely similar in international comparative research reports on computer and internet use, both at school and at home. Table 0.2 in the 2015 OECD report on Students, Computers and Learning (OECD, 2015, p20) shows the ICT equipment and its use at school of the first 38 OECD countries with respect to the number of students per school computer. The Netherlands and Sweden are both listed in the upper half of this list. In comparison with other countries, Sweden and the Netherlands are achieving a reasonable result, but they are also not at the top of the lists of the indicators on ICT equipment and use.

The 2015 OECD report also shows that, overall, in 2012 Sweden was doing slightly better in all the indicators presented on computer ownership and internet usage, but not on the number of computers and computer usage at home. In both countries, virtually all students had a computer at home in 2012, but in Sweden almost 75% of the students had three or more computers at home, versus 69% in the Netherlands. Similarly, Swedish students spend more time using the Internet, both at school (average of 39 versus 26 minutes per day) and outside (144 versus 115 minutes per day). The share of students in Sweden reporting using the Internet for more than six hours per day on a typical weekday is also (considerably) higher than for the Netherlands (13.2 versus 9.9%).

On the other hand, the number of students per computer, and the share of students using a computer at school, is higher in the Netherlands (2.6 versus 3.7 students per computer, and a share of 94 versus 87%, respectively). The share of students using the Internet for schoolwork *at school* is approximately the same, while for internet use outside of school, the shares are slightly higher for the Netherlands.

Other, and more recent data from Skolverket in Sweden regarding lower and middle school students show that between 2012 and 2015, the share of

pupils in grades 1 to 9 who received a personal computer from their school increased from 15% to 30%. In upper secondary school, the share increased from 55 to 80% (Skolverket, 2016).

Both the Netherlands and Sweden were disappointed by both their results and the results in the PISA 2012, as their performance and ranking had decreased since 2009. However, the decrease was noticeably more significant for Sweden and the absolute ranking in 2012 was lower than for the Netherlands. On the other hand, this decrease has continued for the Netherlands in the 2015 rankings, where Sweden has started to perform better once more in comparison with the OECD average, although the performance is still (considerably) lower than in earlier PISA studies and the Netherlands is still performing better than Sweden in 2015. Of course, the question arises as to whether this is somehow attributable to the differences in ICT use in education between the two countries. This question is, however, considerably more easily asked than answered, as the potential relation between ICT use and mathematics and reading performance is merely a correlation and cannot be interpreted as a causal effect.

In the next chapter, several randomized experiments in Dutch secondary education are discussed, from which causal conclusions *can* be drawn.

Chapter 4.
Studies on the Effects of ICT on Education in the Netherlands

This chapter describes several (experimental) studies on the effects of ICT that took place in secondary education in the Netherlands, in most cases at the initiative of the participating school(s), which wanted scientific evidence for the program, teaching approach or hardware that they were using. These studies consist of experiments or quasi-experiments on the effects of intelligent tutoring systems for mathematics and language, such as Mousework and GotIT?!, on the effects of digital differentiation at three levels in a biology class, on the effects of digital feedback and digital testing, and on the effects of using an interactive whiteboard (SMARTboard) in class. All studies were carried out by researchers from the Top Institute for Evidence-Based Education Research (TIER), or by students under supervision of researchers from TIER. The author of this report was involved in all of these studies. Each digital application is studied in one (quasi-) experiment, except for the case of Mousework, which is studied in three consecutive randomized experiments carried out over a timespan of three academic years. In the remainder of this chapter, for each ICT application that is studied, a description of the research is first given, in which the characteristics of the program under study are described, along with the setup of the experiments and a short description of the school(s) under study. Following this, the findings on the effects of each application are described. This chapter concludes with a synopsis in which the findings of each program come together in general findings on the effects of using ICT in education in secondary schools in the Netherlands and discusses the external validation and generalization of these results.

Mousework[7]

This section describes the program under study, Mousework, along with the setup of the experiment and the way the participants are randomized in the study. Furthermore, it explains how student performance is measured and offers a short description of the schools where the experiments took place.

Program Mousework

Mousework is an interactive digital practice tool and a so-called intelligent tutoring system (ITS). The purpose of Mousework is to help students practice their mathematics and language skills, while being able to individualize, and give users direct feedback (Muiswerk, 2013). Mousework modules target pupils from as young as three, when they start to recognize words, up until students in adult education, for example those who have to learn Dutch as a second language. In the Netherlands, around half of the schools use the program in some way, although only a small share of the schools use the program in terms of its original design, namely as a homework tool, next to regular mathematics and Dutch classes.

The schools that are studied offer the program for use at home (via the Internet)[8] and ask students to use the program for half an hour a week per subject, as additional homework (but instead of an extra hour of class, which would have been the alternative). "Mousework" is available for various subjects. The schools under study use ICT for mathematics and Dutch, the mother tongue in the Netherlands. In this chapter, I refer to the "Mousework" ITS as "the tool".

The program is interactive and person-specific. Students work at their own level and access exercises that will help them improve the sub-aspects of mathematics and language in which they are not knowledgeable, while other exercises are intended to maintain their pre-existing knowledge. Students have a certain set of exercises available, covering all domains of mathematics and language, from which they choose when they login into the system. The didactical principles on which the tool builds are differen-

---

7. This section is based on Bartelet et al. (2016); Ghysels and Haelermans (2016); Haelermans and Ghysels (2016); Haelermans & Ghysels, (2017a).

8. An earlier study (Haelermans & Ghysels, 2017a) shows that only few students do not have a computer at home with which to practice.

tiation, making small steps ('scaffolding'), action (practicing) and variation, direct feedback and the teacher as supervisor (Muiswerk Educatief, 2013).[9] The tool consists of tutored exercises, organized into submodules ('sets')[10] and modules which are organized within a specific mathematical domain.

Sets and modules are offered to the student in an adaptive and stepwise manner. At the starting point (typically the start of the year), students take an orientation test which determines, among the various mathematics and language domains of the tool, which sets and modules are offered to the student. At the end of a module, the tool evaluates progress and makes the student repeat (types of) exercise depending on the error level attained in the set. After this repetition within a set, a student can begin a subsequent set within the offered modules.

At regular intervals (intended to be bi-weekly, but in practice once every three to four weeks), students take a short computer test at school to determine in which exercises their skills are lacking and for which exercises their knowledge level is sufficiently high. Teachers are required to organize these tests for their students. After every test, new modules are made available to the student, and the number, type and level of exercises from which a student can choose are adjusted to their new skill level. Without these tests, students can finish the sets pertaining to the modules made available to them and can choose to start over with particular sets, but they cannot make significant progress within a particular mathematical domain, because higher-level content (explanations and exercises) requires them to pass the threshold level at these tests.

It is important to note the balance between choice and structure offered by the tool. Students take a test to gain access to a menu of modules and sets within these modules. They decide in which sets they will participate and are encouraged to finish these sets, because only at the end of a set can they receive a reward (a positive signal) and move to another set within the module. Nonetheless, the fact that students can choose with which set to start also means that students have the option of determining on which mathematical or language domain(s) they will work. The tool does not force

---

9.   Note that I here explain the way the tool is used at the schools under study. Various other "modi operandi" are available (e.g. supervised use at school and non-restricted practicing, i.e. open access to all modules).

10.   A set contains between 8 and 20 exercises, with 12 as the mode.

students to undertake a balanced portfolio of sets. If they do not aim for such a goal themselves, the menu prompts them to do so, but, ultimately, external supervision (typically by the teacher) may be required to ensure it.

The tool's tuition takes place in various forms. Before starting a particular set of exercises, students are invited to read a short explanation concerning the topic of the exercises (a few screens). While performing an exercise, students can ask for help at any time (receive a hint in a pop-up) and after finishing the exercise, the student receives a positive signal when she gets the answer right or more explanation when the answer is wrong. In the latter case, the feedback is as personalized as possible, as the tool contains a database of common errors and the specific feedback these require.

The tool provides exercises in four mathematics domains: numbers, proportions, relations and measurement.[11] For an extensive description of the types of exercises in each domain, please see Bartelet et al. (2016). The tool also provides exercises in six language domains: spelling, vocabulary, grammar, text comprehension, formulating, and listening.

Technically, the tool requires students to provide answers to multiple choice and open questions, but rarely requires them to provide indications of the strategy applied to reach the answer. This is a deliberate choice by the developers, because they want their tool to be useful regardless of the solution method preferred by the teacher (and previous teachers in primary education). In the explanation screens, the most common didactical approach is followed in order to maximize the match between the solution strategy offered at school and by the tool. To facilitate calculations, the tool offers a calculator, which is presented in three forms adapted to the level of difficulty of the exercises.

Previous studies have shown that students are not necessarily intrinsically motivated to do their homework on the digital practice tool, and students use it more frequently when they are motivated to do so by, for example, their teacher or their parents (Haelermans & Ghysels, 2017a). Therefore, Mousework has also developed an app for parents, in response to the belief that parental involvement via an app would increase the amount of home-

---

11. Note that officially the domains are called numbers and operations, ratio and proportion, mathematical relations and measurement and geometry, and the Mousework program has sets and modules accordingly. For reasons of brevity, in this paper we shorten these to numbers, proportions, relations and measurement.

work time students spend on the digital practice tool. This app was free of charge and available for both IOS and Android, and allowed parents to log into the Mousework system with their child's login number (student number). Once logged in, they could see the number of minutes practiced per week, separately for mathematics and language over the course of time. They could also access comparisons between their child's practice behavior and the practice behavior of her classmates, along with comparing their child with him/herself over time. Furthermore, there were performance data available, again over time and compared with classmates, and a suggestion of which aspects of mathematics and language require improvement. Parents could choose to look at numbers or read a short written precis generated from the underlying data. It was possible to add multiple children to the app and follow them simultaneously.

The app registers the child's login number every time the parent logs in. Unfortunately, it was not possible to register what exactly the parent was looking at when logged into the app. In the case of multiple children, the app does register which child the parent has logged in to follow.

The field experiments

Mousework experiment 2012/2013

The first Mousework experiment took place in 2012/2013 among seventh-grade students at one secondary school, among a total of 336 students. In this experiment, randomization took place at the class level. The usual practice at this school is, at the start of the school year, that students are randomly assigned to classes, within the boundary of the ability grouping that forms part of the Dutch system of secondary education ("early tracking") and the option for each student to select two friends with whom to be placed in the same class. Therefore, despite the randomization at the class level, students are still generally randomly assigned to the treatment, as the division over classes was also largely random. The experiment contains only two types of first-year classes (five prevocational classes and eight higher general/pre-university classes). Two classes of each type were assigned to the control group (106 students), with the other nine classes being the treatment group (230 students).

Students in the treatment group were able to practice with the mathematics and language modules of Mousework, and were encouraged to do so by their teachers, whereas the control classes were not allowed to practice

with the tool during the first semester. During the first half of the first semester, students practiced with the tool at home, and had an individualized learning path with selected exercises. In the second half of the first semester, the effects of non-individualized digital practice with the tool were studied, as the school had decided, without deliberation, to give all students access to all possible exercises, regardless of whether those exercises fitted into their individualized learning route. The crucial (and only) difference to the first half of the semester was that the exercises available were no longer individualized. This means that students were no longer offered an exercise program tailored to their skill level, but were instead required to make their own selection out of a much wider range than previously.

Mousework experiment 2013/2014
The second Mousework experiment took place in 2013/2014 among 350 students in 14 seventh-grade classes. Students were individually randomized into treatment and control groups. The intervention in the second experiment is that treatment students practice with the spelling modules of the computerized adaptive practice homework tool. Control students also practice with the online tool, but their accounts are technically disabled for using the spelling modules, and instead they have access to the vocabulary modules. All students have access to all other modules of language, such as text comprehension and grammar. A pilot study with the same program indicated that the online practice tool would potentially only be effective for modules that are easier to automatize and are not so heavily related to the methods used and the pace of the school. It was originally planned to also study the effects of practicing with vocabulary. However, there was such little variation in the outcome measures, that these analyses would not have been reliable. Therefore, the spelling module was chosen. After the first semester, the domains with which the students could practice were switched, to ensure full learning possibilities throughout the school year.

Mousework experiment 2014/2015
The third Mousework experiment, which took place in 2014/2015, actually consisted of two parts, one studying the effects of the Mousework program, and another studying the effects of parental involvement within the setting of the Mousework program. For the part on the effects of Mousework, all students of grades 7, 8 and 9, more than 3000 in total, were individually

randomized into a treatment and a control group. Both groups were to practice with the tool, each with different domains. The treatment group would practice with the modules of only two domains of mathematics, and one domain (the largest) of language, and the control group with the *other* two domains of mathematics and the other three domains of language. In this design, both groups are treated with different modules and are each other's control group. After the first semester, the domains in which the students could practice were switched, in order to ensure full learning possibilities throughout the school year.

In the second part, the effect of parental involvement, via the use of the Mousework app, on students' usage of the app is studied. Parents were individually randomized into treatment and control groups. It is worth noting that this was performed independently from the individual student randomization to study effects of parental involvement regardless of which domains their child could practice. Around half of the parents, the treatment group, could actually login into the app, whereas the other half could not (it was created as much as this was technically impossible). However, *all* parents were asked and encouraged to download the app, in order to receive information on the willingness of parents to use such an app at all, or rather, to gain an idea of the selectivity of parental involvement using a digital tool such as this. Only after downloading and logging in (or trying to) would parents find out whether they belonged to the treatment or control group. Parents were pre-informed about the experiment in both an information letter and during annual parent information meetings at the start of the school year. Parents who belonged to the control group would, upon trying to log in, be given a message reminding them of the experiment and clarifying that they would be able to login into the app after the current semester (i.e. for the second part of the schoolyear).

Although treated parents could use the app as often as they desired, they were advised and requested to use it at least once a week.

Student performance

The mathematics and language skills, which are the performance measures in this study, are measured using digital standardized mathematics and language tests, which are taken by all students that participated in one of the three experiments, at the start of the school year, and after the first semester. These are standardized validated tests developed by the company behind

the tool, and these tests are based on other nationally validated tests. The reliability (Cronbach's alpha scores of between .79 and .92) and validity of these tests are analyzed annually by the tool developer, based on norm data of several participating schools (Schijf & Schijf, 2014). Although the pre- and post-tests are digital tests that are developed by the same company as the tool and are administered in the same digital environment, the tests themselves are external to the practice exercise tool and do not contain any of the exercise questions. The tests measure whether students have mastered the required national mathematics and language level (called 'reference level'), given their age and the fact that they have finished primary school.

The mathematics test consists of relatively simple multiplication or addition questions, but also contains special understanding questions, where the student sees an unfolded shape and is asked to select the figure that could create it. Alternatively, the student is asked to calculate the volume of a sphere, or to quickly make mental calculations. The test contains multiple choice questions and students were allowed to use scrap paper for their calculations, but not digital calculators. The tests last around 20 minutes. The language tests, for example, consist of spelling questions, vocabulary questions, text comprehension, grammar questions, and having to listen to some information and then answer a related question. The test lasts around 90 minutes. The score of both the pre- and post-tests is recalculated (by the tool's developer) to a score between 0 and 100.

The schools under study

The first two experiments are carried out at one secondary school, the third experiment is carried out at three secondary schools, including the school that also participated in the first two experiments. The three schools under study, all located in the Southern part of the Netherlands in the province of Limburg, have between 1500 and 2500 students, and are, by Dutch standards, mid-sized schools for secondary education (junior high and high school). All three schools offer secondary education in all tracks[12] and are tracking students from the seventh grade onwards in several prevocational, general and pre-university tracks. Compared with the average Dutch secondary school, these schools have a higher graduation percentage (around 95% ver-

---

12. See Chapter 3 for more information on the offered tracks in secondary education in the Netherlands.

sus 92.5%). All three schools are doing relatively well performance-wise. For the first two experiments, all students in seventh grade at the one school were included in the research. In the third experiment, all students in grades 7, 8 and 9 for all three schools were included.

Findings

In this section, I describe the findings from the three Mousework experiments. First, the general effects, principally stemming from the first Mousework experiment, are discussed, followed by the differential effects for both student achievement groups and different domains of mathematics and language, stemming from all three experiments. Next, the results of student practice behavior and the teacher role are discussed, both based on all three experiments. Lastly, the findings on the effect of parental involvement are discussed, based on the third Mousework experiment. This section concludes with some general conclusions on all three Mousework experiments.

General effects

The overall results of the three Mousework experiments show that, in general, there is a positive effect of practicing with the Mousework program, but the conclusion is also that this primarily holds for the mathematics domains and the easier to automate domains of language, such as spelling and grammar. No effects were found for language domains such as listening, text comprehension and formulating.

The first Mousework experiment is the only one that studies the overall effect of Mousework, without dividing mathematics and language into their separate domains. In this experiment, the control condition does not practice with Mousework at all. In this first experiment, we focused on the effects of an individualized interactive digital practice tool on general mathematics and language skills for seventh-grade students. The results from the first part of the experiment show that there is a statistically significant effect of (being able to) practice individualized with the digital tool on both the absolute mathematics score and the growth in the mathematics score, as well as on the share of students increasing in the reference level for mathematics. The effect on the growth in score is 4.8 points, on a scale from 0 to 100, which corresponds to a small to medium effect of 0.40 of a standard deviation. Given that the increase in mathematics homework time

due to the tool was 25 to 33% (an additional 15 minutes), this small to medium effect is not surprising. However, this effect might not solely be due to practicing, but could partly be caused by the small tests that were given three or four times at school, during the first part of the experiment (relates to the testing effect, where usually large effects are found, see Roediger & Karpicke, 2006), and possibly also to the increased teacher attention for the treatment classes. Furthermore, there was no effect at all for language performance during this first part of the first experiment.

In the second part of the first experiment, the treatment group did not have an individualized learning path in the Mousework tool, but instead had access to all existing mathematics and language exercises. The control group did not practice with the tool, and also did not have additional instruction or homework time. The results from the second part of the first experiment show no effects, although students practiced, on average, more minutes per week during this period, implying that the effectiveness of the tool found in the first part of the first experiment is primarily due to its individualized and personal nature (similar to the findings in Barrow et al., 2009), and not merely to the additional practice time. Again, no effect was found for language.

The results from the first Mousework experiment led to a dual conclusion: 1) Individualized digital practice tools are effective at improving the mathematics performance of secondary students, and 2) it is the individualization of exercises that makes digital practice tools effective, and not the additional practice time in itself.

Differential effects

In all three Mousework experiments, differential effects among different domains and achievement levels were studied. In the first experiment, the control condition was, as mentioned, not practiced. In the second and third Mousework experiments, the effectiveness of practicing with specific mathematics and language domains was studied, where other domains served as the control condition. Therefore, all students practiced with Mousework, just not with the same domains. In the second experiment, the effect of practicing with the domain spelling was studied, in the third experiment pretty much all domains for both mathematics and language were studied.

From the first experiment, the conclusion is that there is a differential effect, both for achievement groups and for the mathematics domains. Al-

though the implementation of a non-compulsory supplementary tool leads to improvements in the overall mathematics skills of both low- and middle-achieving students, but not for high achievers, it is significantly less effective for middle achievers than for low achievers. This finding is especially true for the analyses divided among mathematics skills domains (numbers, proportions, and measurement). It is also found that having access to the tool has, on average, a statistically significant positive effect on students' proportions skills, but not on their numbers and measurement skills. Thus, on average, a differential effect of the tool is observed across distinct mathematics domains, although this also differs for students of different achievement levels. Low-achieving students benefit significantly from the intervention in all mathematics domains, while middle-achieving students benefit significantly from the intervention in the domains of proportions and measurement.

In the second experiment, the outcomes of an ICT experiment with the aim of improving spelling skills among seventh-grade students were studied. Since students were intended to use the tool as additional homework, it is not surprising that the adoption of the tool by students is uneven. Furthermore, among teachers, the enthusiasm is mixed, in part because the freedom it offers to students makes it difficult to coordinate home practice learning with a regular, class-based and fairly homogeneous instruction process, as will be explained in the following paragraphs.

Nevertheless, the experiment reveals a positive effect of giving students access to the tool (without taking into account how much they have used it), suggesting that uneven student adoption and teacher support did not hamper the average benefit to students. Moreover, an analysis in which the actual usage of the tool by students is taken into account also shows positive effects. Using the tool for spelling practice substantially contributes to the spelling performance, although the analyses reveal that effects are solely present for low-ability students, with an effect size of 0.16 of a standard deviation. Therefore, getting more students to use the tool, or to use it more actively, may lead to stronger results than observed in this trial. Furthermore, it might be the case that students would also have gained if they had used non-digital methods to practice spelling, which cannot be ruled out, and, unfortunately, also cannot be tested with this study design.

These findings are based on whether the students use the tool at least once. However, if we look at the intensity of treatment, we learn that the

effect primarily appears to concern using it at all. With the average number of times and minutes per week that we observe, we see a positive statistically significant effect. However, students that *do* practice would need to double the amount of minutes per week to achieve a sizeable effect for additional practicing. Therefore, the effect that we find on this 16-week experiment seems to be small.

In the third experiment, the effect of the four separate mathematics domains and the four out of the six separate language domains were studied for the three grade levels and the different levels of education separately. Several differential effects were found, but barely any overall effects. This is most likely due to the similar control conditions, as the control group also practiced with Mousework for both mathematics and language, just with different domains. As for the mathematics domains, all four were effective, but not all for the same group of students. Practicing with numbers and relations has a positive effect on the mathematics performance of seventh- and eighth-grade students. Practicing with proportion and measurement, on the other hand, primarily seems beneficial for the better performing eighth grade and all ninth-grade students. It seems that the easier domains are more beneficial for lower graded and lower performing students, whereas the more challenging domains are more beneficial for better performing and higher grade students.

As regards language, only an effect for spelling in the third experiment was found, an effect that was also already found in the second experiment. The spelling effect was found for seventh- and eighth-grade students. Furthermore, a positive effect for grammar was found, but only for seventh-grade students. No effect at all was found for text comprehension and formulating, and also no effect for language for ninth-grade students at all.

Practice behavior

In all three experiments, student practice behavior was a source of concern, as students in general practiced significantly less than the 30 minutes per week for mathematics and the 30 minutes per week for language that were asked of them. In the best-case scenario, they only practiced around half of those 30 minutes per week. However, more positively, it was still possible to detect the positive effects of practicing with the tool, despite the fact that students did not practice that often. In all three experiments, the practicing behavior of the students was generally related to the teacher, and, to a much

lower extent, also to some student characteristics. Higher-performing students, for example, practiced more.

Aside from the abovementioned general conclusions with respect to practice behavior, this was also studied in more detail in the first experiment, where the focus was on explaining students' practice behavior for mathematics. As explained above, students are intended to practice at home with the tool, and to regularly complete small tests at school. If students do not take tests, they will not have new exercises with which to practice. There are two factors that determine whether students write these tests and how many. First, the teacher has to take the class to a computer room to enable students to write the test. Second, students need to have practiced as well, in order to have a test available to take. Whether a student has completed any small tests is therefore dependent on factors that the student can influence (practice) but also by factors that the student cannot influence (teacher organizing a computer class).

The results of the descriptive analyses of students' practice behavior show that, especially for the domains of proportions and measurement, only a few students have actually practiced. Overall, the students from the middle skills level group practiced the least. However, separately, the practice data for the domains show that the middle group students practice the least in the domain measurement, and that low-skilled students practice the least in the domain proportions and high-skilled students in the domain numbers. Furthermore, the results show that practice and test behavior are different among skills level groups and domains, but also within skills level groups and domains there are large differences between students. The number of students that practiced for the domains proportions and measurement are very different than for the domain numbers. This implies that only a select group of students have practiced in the latter two domains.

Teacher role

In the light of student practice behavior, it is found that a teacher who uses the tool more, and a teacher who administers more small tests at school, have students that practice more. However, a teacher who simply has a more positive attitude to the tool has students who practice significantly less. The latter may indicate that it may not be sufficient to have a positive attitude, but an active interest in the tool may also be required to motivate and advise students on the use of the tool.

Therefore, it may be clear that teachers play a large role in the program Mousework, as it is generally teachers that create the student practice. On the other hand, despite low teacher enthusiasm and lower teacher time investment, the experiments still show that Mousework is an effective program to increase mathematics and language (spelling and grammar) performance. Furthermore, the program works somewhat independently from teaching practices in class, so, from that perspective, the teacher's role is not necessarily extensive. However, as the teacher is key for getting children to practice, the potential gains of the program are considerably larger than what has been shown so far. On the other hand, it does not necessarily have to be the teacher who stimulates the students to practice; this could, for example also be their parents, or their peers by means of a competitive element in the program. This competitive element was not part of the Mousework program during the three experiments, but the role of parental involvement was studied. This will be discussed in the next section.

Parental involvement

In the study on parental involvement, the effects of parental involvement on the use of a digital homework practice tool and on the mathematics performance of all students in grades 7 to 9 of two secondary schools in the Netherlands were analyzed. The experiment consisted of an app in which parents can follow their child's practice behavior in the digital homework tool, using a randomized field experiment at the individual level. For additional information on parental involvement, both students and parents were asked to fill out a questionnaire.

The results were analyzed from both the viewpoint of the provision of access to the tool and concerning the intensity of the effective use of the app, controlling for non-compliance of parents who did not use the app, even though they were allowed to do so. Both analyses show that parental involvement via app use positively affects the practice behavior of seventh- and eighth-grade students, but negatively affects the practice behavior of ninth-grade students. Furthermore, positive effects of the use of the app on students' mathematics score at the end of the experiment are found, which is primarily driven by the eighth-grade students.

Subgroup analyses show that the positive and statistically significant effects that are found (both on the use of the homework tool for grades 7 and 8 and on mathematics performance in grade 8) are due to the male students

and the low-SES students, whereas the negative effects of the parental app on the use of the homework tool in grade 9 are principally due to the high-SES students.

As such, these results add to the somewhat limited existing experimental literature on interventions to raise parental involvement. In contrast with intensive interventions such as those reported by Bergman (2015) and Mayer et al. (2015), providing access to a parent app linked to an existing digital homework tool requires little effort for the school and the teachers. Nevertheless, it has proved effective in raising involvement and is beneficial to the learning progress of the students.

Both the parental and the student questionnaire shed additional light on how students and parents experience parental involvement and the students' feelings in this regard. For seventh- and eighth-grade students, parents and students are very much aligned with respect to how they feel about parental involvement and when it might be needed, whereas there is a clear discrepancy in this regarding the ninth-grade students and their parents.

Cost-effectiveness

Determining the cost effectiveness of this digital practice tool can be performed from the schools' point of view, but also from that of society. The costs of the digital practice tool for numeracy or literacy separately are approximately 25 euro per student per year, and, for both of the packages combined, around 40 euro per student per year. Of course, this only includes the cost of the program and not the cost of having to purchase a tablet or computer, if necessary. For the students in our study this was not relevant, as they all had access to a digital device of some sort. The total cost for this tool for a group of 400 seventh-grade students for a school is around 10 000 euro. For the school where we conducted experiments three consecutive years, an alternative measure to foster numeracy skills could be the introduction of an additional mathematics class (for the school under study, this was the alternative they considered). The additional costs of hiring a teacher who practices numeracy skills with the student for at least an hour a week would be considerably higher than using this digital tool. Given that there are 13 first-year classes, one would need an additional full-time teacher each year, entailing annual costs of at least 30 000 euro.

With respect to the gains for society, we see that around 85% of the first-year students in our dataset of the first experiment performed lower on

numeracy than their predicted results according to the national reference levels (Commissie Meijerink, 2008), having finished primary education successfully. If practicing with the digital tool were to increase the average numeracy level such that the majority of these students would perform at the expected level in their graduation year, the financial savings to society could be considerable. Each student that does not fail the national exam at the end of secondary education because of failing their numeracy exam saves the government 7 000 euro, which is the average cost per student per year (Teule, 2012). Furthermore, retentions in grade will delay the student for at least one year in entering either vocational education or higher education, which, in turn, delays her entry into the labor market by at least one year. The opportunity costs of the student will therefore be considerably higher than the 7 000 euro for the government. In any case, to be cost-effective from society's point of view, the introduction of "Mousework" in the starting year of secondary education only needs to allow two students out of 400 to graduate on time instead of delaying their graduation with one year. Given that our study has shown that it can be expected to help 10% of the students across the threshold (Haelermans & Ghysels, 2017a), the latter seems highly likely, although future research following up on students throughout secondary education should confirm that expectation before any concrete statements of this type can be made.

Conclusions: Mousework experiments

Based on three years of experimental research related to the Mousework program, it can be concluded that, in general, an interactive practice tool, or an intelligent tutoring system, such as Mousework, is beneficial to students for mathematics and for the easier to automatize domain of language (e.g. spelling and grammar).

However, some domains are more beneficial for some types of student than others. The easier domains of mathematics, spelling and grammar are primarily beneficial for students from grades 7 and 8. The more difficult domains for mathematics are more beneficial for higher performing students and students in grades 8 and 9.

Three years of experiments also show the large diversity in practice behavior by students. On average, students practiced considerably less than they were intended to, only half of the intended 30 minutes per week, and in many cases even less. Furthermore, we see large differences in practice

behavior between the different domains of mathematics and language. Although student characteristics such as previous performance level and age can explain differences in practice behavior to some extent, the analyses reveal that it is generally the teachers, and whether they encourage practicing and organize small tests at school, that make students practice. Furthermore, stimulating parents also makes a student practice more frequently.

Despite the fact that the teacher is the largest factor in the practice behavior of the student, the effect of Mousework remains, even if the teacher is not enthusiastic and motivating. Furthermore, the program also functions by itself, and it is not necessarily related to teaching practices in class, thus giving the teacher a dual role. The potential benefits of Mousework could be much greater if the teacher is actively involved, or uses the results from Mousework to focus more on certain aspects in class (using individual or class dashboard on how students perform in certain topics). However, it also works without the teacher.

Aside from the teacher, parents can also encourage practice behavior and, in this manner, increase student mathematics performance. If parents are more involved via the Mousework app, seventh- and eighth-grade students practice more and perform better. However, the reverse is true for ninth-grade students, as they practice *less* if their parents are too heavily involved. Furthermore, the positive results of seventh- and eighth-grade students seem primarily attributable to low-SES (socio-economic status) students.

Overall, it can be concluded that a digital interactive practice program/intelligent tutoring system such as Mousework is very promising for practicing basic skills such as mathematics and language, as long as the skill can easily be automated. However, students do need motivation to practice, and, for this, other actors are needed, such as teachers or parents. Moreover, although there are differences in practice behavior and motivation between, for example, boys and girls (the latter practicing more) and different grade levels (students in lower grades practice more) we do not find any difference in the effect on general performance of gender, grade level, or socio-economic background.

GotIT?[13]

This section describes the program under study, GotIT?!, along with the setup of the quasi-experiment and the way in which the participants are selected for the study. Furthermore, it explains how student performance is measured, and provides a short description of the schools where the experiments took place.

Level differentiation via the program GotIT?![14]

This section considers a Dutch computer-assisted online tool (CAI) called GotIT?!. The hallmark of this educational software is that it was created through consideration of the best approaches to teaching mathematics as well as the needs (cognitive, psychological, etc.) of the students. The CAI tool offers a large number of exercises at different difficulty levels. This enables each pupil to organize the work and progress at a rate consistent with his/her own level of ability. This allows the teacher to differentiate within the class. Pupils who experience fewer difficulties with the theory and advance quickly in solving exercises can move on to exercises of a higher level of difficulty without being impeded by pupils who progress more slowly. As the tool is easy to master, it is unlikely that there is a significant difference among teachers in the mastery of the tool. The CAI tool is adaptive in that it adjusts its exercises to the knowledge and level of the student. GotIT?! provides pupils with tips on organization and skills for solving exercises, all of which may benefit the pupils' confidence in the learning content, improve their metacognitive skills and provide a way for skill drills (i.e., practicing an activity until it becomes automatic). The content is organized around 11 subjects. These include, for example, addition, multiplication or counting principles.

Moreover, the GotIT?! system offers features that provide prompt and continuous feedback to the teacher on pupils' learning progress both at the level of the individual pupil as well as the classroom. More precisely, the feedback and control system comprises tools for tracking each individual pupil's step-by-step progress so that, at each moment, an accurate overview of his/her competence level is possible. In this way, the teacher can monitor

---

13. This section is based on De Witte et al. (2015) and an unpublished paper on the same topic.
14. Description of program GotIT?! taken from De Witte et al. (2015).

which pupils realize the milestones and which require additional attention. Based on this continuous stream of information on pupils' progress, the teacher can determine whether an adjustment in the instruction approach or any other type of remediation is warranted for the class as a whole or for one or more individual pupils. In addition, GotIT?! also includes communication features that enable teachers to interact and communicate with the pupils both at a classroom level as well as individually. Depending on the circumstances, the teacher can decide to provide feedback to all pupils in the class, a subgroup of pupils or just one individual pupil.

The quasi-experiment

In collaboration with the software designer of GotIT?!, a list was created of schools that intensively use the software in their classes in the seventh, eighth and ninth grades of secondary school in the year before the experiment. In an attempt to avoid selection bias, matches to these schools were sought, based on comparable characteristics on indicators such as the school type, size, number of teachers, and percentage of students from disadvantaged areas, as defined by Statistics Netherlands (CBS). While the matched schools have similar observed characteristics, the main difference between the two groups is that they use or do not use adaptive ICT tools during mathematics classes. The data for the observed characteristics are drawn from publicly available data from the Dutch Ministry of Education, Culture and Science. The assumption is that, if the schools are similar in regard to the observed characteristics, they are also similar in regard to the unobserved ones. Schools which use the adaptive ICT tool are labeled as 'ICT schools'. The other half of the schools on the list consist of schools which have chosen to teach mathematics using traditional and non-ICT-related tools (i.e., the traditional way). These schools are labeled as traditional schools.

The selection of the schools took place in September 2013. From November 2013 to December 2013, all schools were contacted and asked to participate in the research. Primarily due to time constraints, more than half of the schools that were asked to take part declined to participate in the research. Ultimately, there were five 'ICT schools' and two traditional schools that participated in the research, with a varying number of classes, but with a minimum of three classes per school.

The pre-tests took place in the last week of January and the first week of

February 2014. The test consisted of a standardized 'VAS' test, developed by the independent testing organization, CITO. While a different test for each age group was provided, the educational levels (prevocational education, higher general education and pre-university education) were mixed in the test. The test was electronically distributed through Google Forms. Around 836 students took this pre-test. An analysis of the schools based on known characteristics as well as pre-test scores shows that they are comparable in these characteristics. The post-test was taken in a similar way to the pre-test in the last two weeks of May 2014, by about 400 students. It also consisted of a standardized 'VAS' test by the testing organization CITO.

In a second part of the research, a regression is estimated that correlates the intensity use and school attainments for the ICT schools that only use GotIT?!. It is worth noting that this analysis is not causal, but can give some interesting insights into what drives the effect found in the first part of the analysis. The intensity of working with the CAI tool is proxied in two ways. First, the number of exercises a student has made is considered: the greater the exercises, the more intensely the student has worked with the tool. Second, the number of subjects the student has successfully completed is considered. By combining the two outcome variables, complementary information is obtained.

Student motivation and performance

Together with the pre-test, all students were asked to fill out a detailed questionnaire on their use of and experiences with ICT in general, at school and during mathematics classes. Through 31 different questions, they were asked about their motivation for their education in general and for mathematics specifically. The questionnaire is inspired by Van Braak et al. (2010), in order that it can be internationally compared and has previously been validated. The questions included: 'I like to attend the mathematics courses', 'I think that mathematics is useful outside school', 'I can pay better attention to the class when using computers', 'I like learning more when using computers', 'Computers make courses at school more interesting', 'I use the computer for YouTube', 'I use the computer for gaming', etc. All questions are on a Likert scale between 0 (no agreement; never) to 5 (total agreement; daily). The questions show a high internal consistency, as will be shown later in more detail.

The pre- and post-tests were designed in such a way that they took

around 12 minutes to fill out. Of the respondents, 17 students were removed from the data who did not entirely/appropriately fill out their name or class. This was considered as a sign that the students would also not respond seriously to the actual test questions, and would thus bias the results. Furthermore, only students that participated in both tests were included in the analysis. This excluded more than half of the students as, during the research, some teachers refused further participation as they considered filling out the test as too time-intensive for their students. A careful examination of the excluded students indicates that this is not necessarily a random sample. The students that were excluded, for example, significantly differed in the average level of education, gender (more male students were excluded), their motivation for mathematics, and their year of study. However, it seems as if these students that were excluded came equally from both the treatment and the control group, as the remaining treatment and control students are still largely comparable.

The schools under study

The schools that use ICT intensively apply the adaptive computer program 'GotIT?!' for their mathematics courses (www.got-it.nl). The program provides pupils with suggestions on organization and skills for solving exercises. It is intended to improve students' meta-cognitive skills and allows students to intensively practice mathematics exercises. As a key feature of the program, students can create the exercises based on their own abilities and pace.

The decision to teach mathematics courses with ICT or in a traditional manner is, in these schools, taken by the coordinator of the mathematics courses or the school management, following deliberation with the mathematics teachers. In order to secure continuous learning styles across grades and classes, in most schools, all mathematics teachers within the same school use the same didactical tools.

Findings

The results from the first part of the quasi-experiment indicate that seventh-grade students in ICT schools have a higher growth path in learning outcomes than students in traditional schools. A relatively large standardized effect of more than 0.6 of a standard deviation is found for seventh

grade students. However, the effect is not statistically significant for students in eighth and ninth grade. It therefore seems that the effect of ICT on learning outcomes is different for different age groups. While ICT has larger merits in the first year, in later years it does not matter whether ICT or traditional teaching methods were used. Another potential reason is that seventh-grade students were new to the school (coming from primary education) and the use of 'GotIT?!' was new to them, actually making a difference, whereas the eighth- and ninth grade-students had already been using 'GotIT?!'. The finding that there is an effect for seventh grade students is similar to the results found in the previous section on the effects of Mousework.

The results from the second part of the study, which are not causal (so we cannot draw causal policy conclusions from this), can give some insight into what drives the effects that are found in the first part. This correlational analysis shows that making more exercises is correlated with higher test scores.

To summarize, it is found that seventh-grade students, in particular, benefit from having access to and practicing with GotIT?!. Furthermore, it is found that, given the participation in the CAI tool, taking more exercises is associated with higher test results for all students.

SMARTboard[15]

This section describes the SMARTboard intervention, along with the setup of the experiment and the way the participants are randomized in the study. Furthermore, it contains a brief description of the school where the experiment took place.

In-class-level differentation with the SMARTboard

The intervention consisted of in-class-level differentiation, made possible by effectively teaching with an interactive whiteboard, called SMARTboard. The SMARTboard was used to differentiate learning materials among

15.   This section is based on Cabus et al. (2017).

students, communicate more directly with students, provide students with additional learning materials when relevant and providing students with more independence in their learning process. Therefore, the primary focus of teaching with a SMARTboard was level differentiation between the students in class, as well as the speed at which a student can understand and perform mathematics exercises in class. Treated students received 50 minutes of SMARTboard education in basic mathematics skills each week over a period of six weeks between autumn and the Christmas break of 2013, whereas control students also received basic mathematics skills instruction, but without the SMARTboard. Both the control group and the treatment group followed the same instruction book for mathematics. However, along with the instruction book taught using the SMARTboard, the treated students received additional (digital) instruction material, such as education on the basis of notebook software, and websites especially designed for teaching mathematics with the SMARTboard on the Internet. They did not receive homework, but instead students could practice in small (same level) groups in class, during instruction time, to work on their mathematics assignments. In case they finished their mathematics assignment earlier than the rest of the group, they could participate in another group in class working on another, more difficult, task. Due to the use of the SMARTboard, the traditional class teaching time of the teacher was reduced, so that the teacher could explain certain aspect in the level groups, but students could also explain to each other. Students performing at a higher level could often work without teacher explanation, allowing more time for the lower achievement groups for the teacher. Furthermore, for the students, more time was available in class for the task, which would otherwise be homework. The underlying pedagogical approach was, therefore, in-class-level differentiation, both at the pace and level of instruction and exercises, and made possible through the use of the SMARTboard technology. Without the SMARTboard technology, and given the current teaching settings (i.e. one class of about 30 students with one teacher), in-class-level differentiation would not have proven possible. In contrast to the intervention group, the control group only received classical instruction in class, and regular homework. At the end of their 50 minutes of instruction time, they received homework tasks as usual to be individually undertaken outside school. As such, nothing changed for students in the control group, compared with students in earlier cohorts.

It is strongly believed in the literature that ICT devices can only increase student performance when teachers know how to effectively use them (e.g. Abbitt, 2011; Koehler et al., 2007). In order to meet this necessary requirement, teachers were selected for this experiment based on the three criteria that assess their TPACK (Technological Pedagogical Content Knowledge, a framework that emphasizes that the teacher needs knowledge at pedagogy and contents (expertise), but also at [how to use] educational technology): (1) they had knowledge of ICT in general and of how to use ICT in the classroom; (2) they participated in the course "training-the-teacher using SMARTboard", and (3) they had content knowledge of (teaching) mathematics. The TPACK teacher training consisted of three sessions of six hours each, in which teachers were trained to combine content, didactics and ICT using SMARTboard. This included developing interactive lessons with the SMARTboard, as well as practical knowledge of the SMARTboard, for example how to create links, how to create moving items, and how to use short movies. The last session also included aspects where these teachers were taught to teach their colleagues how to use the SMARTboard in class. The teachers of treated and control classes were, on average, the same, except that the treated students only have mathematics teachers who received ICT training (i.e. fulfil requirement 2), while control students only have mathematics teachers that did not (currently) have the ICT training.

The field experiment at the school under study

The intervention took place in the seventh grade at one average sized secondary school in the Southern part of the Netherlands. Whereas the national standardized exam results at the end of primary education are used to track students into school types, this school uses didactic age equivalence (DAE) to assign students to classes at the start of secondary education. The DAE denotes the total number of months of primary education a student has followed given his/her educational proficiency level. Therefore, there is a list with the expected performance level of a child when she has spent a certain amount of months in education. The list starts at grade 1 (around six years old), and assumes 10 months of education each year. Following this, a test determines the actual performance level of a child, which is then compared with the list, in order to determine the DAE for that child. This can diverge from the actual months in education, as not all students perform in accordance with the national standards/expectations. For example, if, at

the end of primary education, a student has a DAE equal to 40, then this student is performing at the level of someone who has had 40 months of education in the primary school, even though this student has been attending school for 60 months (6 years, each of 10 months). As such, this student has a delay of (60-40) 20 months at the start of secondary education. In conclusion, if the total number of months is lower than the expected value, then the student has a somewhat low prior mathematics proficiency.

As a result of this school policy, homogeneity is obtained within classes and heterogeneity between classes with respect to average mathematics performance. Based on average DAE, the school distinguishes between three levels: A, B and C, with A consisting of the worst performers, and C the best.

Of course, the national and school policies have consequences for the assignment rule to either intervention group or control group. Therefore, stratification is used with respect to the three levels, A, B, and C, before randomly assigning classes to the intervention group. With respect to level A, two of the four classes were randomly selected into the intervention group, while for level B one of the three classes was selected, and for level C, again two of the four classes were randomly selected. As such, there were 11 different classes in total in the seventh grade, of which five participated in the intervention, and six were control classes. Analyses based on observable characteristics show that the treatment and control group were comparable in these characteristics.

Findings

The results show that teaching with SMARTboard is an effective method in order to increase students' mathematics proficiency for treated students. The results indicate a standardized effect size of 0.11. This experiment shows that the performances of students are significantly better for (low-performing) students in classes with SMARTboards.

As additional analyses, the teachers from the intervention group were asked why teaching using SMARTboard in class has positive effects on mathematics performance from their point of view. They argued that students better understood the study materials by graphically plotting the answers on the screen. Because of time constraints and the different mathematics levels of the students, they did not have time to visualize the

answers on the school board, retain all students' attention or help them with problems at the same time. As such, the teachers said they could use the SMARTboard as a tool that facilitated their didactic approach and, at the same time, supported their class managerial skills. One teacher also pointed to the 'competition mechanism' between the groups of students formed in class based on their mathematics ability. In particular, the fact that the solution to their questions was put on the SMARTboard screen boosted the competition between the groups, as they all struggled to find the correct answers. Other contributions of SMARTboard teaching mentioned referred to the extra personal time spent with low mathematics performers. While more able students from the intervention group who finished their exercises early received additional tasks, lower ability groups solved the questions at their own pace and, when needed, with help of the teacher.

Digital differentiation, feedback and weekly testing[16]

Description of the research

This section describes the experiments on digital differentiation, digital feedback and digital weekly testing in eighth-grade biology classes, and describes the setup of the experiments and the way the participants are randomized in the study. Furthermore, it contains a short description of the school where the experiments took place.

Digital differentiation, feedback and weekly testing

In all three experiments, biology classes were studied in eighth-grade (second year of secondary education in the Netherlands) prevocational students (the lowest track in the Dutch tracking system). In the first experiment, all students completed small weekly formative tests that covered the material of the previous week (around 10 minutes, on the computer). The treatment consisted of a differentiated learning process, with content levels adjusted weekly to the student's performance on a small weekly test, whereas the

16.   This section is based on Ghysels et al. (2014); Haelermans et al. (2015a); Haelermans et al. (2015b).

control condition was at a constant level, in accordance with the track of the student, maintained throughout the experimental period. The levels offered were modeled according to three different tracks in Dutch secondary education: the practical prevocational track, the theoretical prevocational track and the higher general track. It is worth noting that it was possible for treatment group students to follow a different track each week, depending on their score on the test. The lessons at the practical prevocational track level were written in more simple language, using fewer words and less complicated sentences. Furthermore, the pace of the exercises would be somewhat slower, which means these students studied the minimum amount of topics. The content of the basic topics did not differ between the groups. The theoretical prevocational track entailed some extra topics compared with the practical track and more difficult explanations and more challenging exercises. This was also the case for the higher general track, where more topics were discussed (in most cases slightly off topic, but interesting to the students). Control students were always offered learning content on the intermediate level, the theoretical track.

In the second experiment, all students took small weekly formative tests, and the students in the treatment group received specific feedback, depending on their answer, after each question. Feedback was provided on each question answered by students. The feedback focused on whether and why a certain answer was wrong, if the question was answered incorrectly, but was also provided to clarify why this was indeed the answer, if the question was answered correctly. At the end of the formative test, treatment group students were provided with an overview of which questions they answered correctly and which questions they answered wrongly. Control group students did not receive specific feedback on their answers during the test, and they were only able to see which questions they answered correctly and which questions they got wrong at the end of the course.

In the third experiment, the treatment group received the small weekly formative tests, whereas the control group did *not*. At the end of this weekly test, treatment students were provided with an overview of which questions they answered correctly and which questions they answered wrong, but without further feedback. In the meantime, the control group students worked in their own digitalized environment on the topic of that week and/or created (digital) homework assignments.

During the school year, the students had two biology lessons (50 minutes

per lesson) each week. During the first experiment, both these classes were computerized, and all students used the digital instruction material, which was a combination of the digital instruction package from the book publisher, sections from the book that were digitalized by the teacher and additional digital material from the internet (such as exercises, puzzles, short movies, etc.).

During the two other experiments, one of the weekly lessons relied on computer-driven instruction and took place in a computer room, whereas the other was teacher-directed and took place in a regular classroom, with some occasional group work. All students used digital instruction and assignments, as well as their workbook, together with classroom instruction, in order to study the course content. The digital instruction consists of the same combination of materials as in the first experiment.

The field experiment

Randomization took place at the beginning of the school year 2012/2013. Stratified randomization at the student level was applied, in order to ensure that the experiment group and the control group were equally distributed in regard to the students' primary school ability test score, the class to which they belong, their gender and age, and state of grade repetition. In effect, a simple stratification algorithm was implemented, allocating students alternately to the intervention or the control group after a random start using a list of students arranged on the previously indicated characteristics within each of the five participating classes. From the total of 115 students, 57 students were assigned to the treatment group and 58 to the control group. All three experiments had the same treated and control students.

The first experiment consisted of 12 weeks in total, distributed over three topics, each of four weeks. It is worth noting that the treatment, which will be discussed later, only affected the content that is discussed, not the skills needed to process the content or other aspects that could influence the pre-test of the next topic because of changed study behavior. The first topic was "metabolism and respiration", the second was "blood circulation", and the third was "your health". Because the school desired students to be graded separately for all topics, there is a pre- and post-test for each of the three topics.

The second experiment had a duration of eight weeks in the spring of 2013. The topic of this second experiment was 'sexuality and relationships'.

The third experiment, information provision via weekly formative tests, consisted of six weeks in total, in the summer 2013. The topic of the course taught during these six weeks was 'heredity and evolution'.

It should be noted that the cumulative nature of the experimental process requires careful attention in any subsequent analyses. Students were exposed to three experiments consecutively and hence it cannot be excluded that some transfer may have occurred. However, none of the statistical analyses conducted so far (Haelermans et al., 2015a; 2015b ; Ghysels et al., 2014), have revealed any indication of transfer. At the starting point of any experiment, the pre-test scores of the experiment and control groups were very similar. The latter suggests that the benefit of the experiments was closely linked to the topic and/or dissipated rapidly.

Student motivation and performance

The pre- and post-test consisted of questions on the covered topic that originated from a test databank with questions that have been used by this school for biology for several years. Test items are therefore not scientifically validated, but are selected on face validity by the various biology teachers which have been involved in the course in recent years.

All tests are computerized, which facilitates the subsequent use of test results in the experiment. The pre-test was multiple choice and took around 15 minutes. It was intended to reflect prior knowledge students may have on the topic to be introduced.

The small weekly (formative) tests also consisted of entirely multiple choice questions, lasting about 10 minutes. The tests check topical knowledge students should have acquired in the two previous classes and the corresponding homework.

Finally, the post-test (of every experiment) was written during one biology class, and lasted around 50 minutes. In contrast to all other tests, the post-tests included both multiple choice questions (around 75% of the test) and open questions (around 25% of the test). The tests probe for knowledge and understanding of the topic of the teaching period that the test concludes.

All pre- and post-tests, as well as the formative tests, are scored from 0 to 100, representing the percentage of questions answered correctly.

Motivation is measured based on Boekaerts' validated online motivation questionnaire (2002). This questionnaire consists of 25 questions that have

to be answered based on the 4-point Likert scale, ranging from not agree at all (1) to very much agree (4). The answers to these questions can be reduced to six components, namely self-efficacy (SEFF), success expectancy (SE), task attraction (ATTR), perceived utility (PU), task anxiety (TA) and intended effort (IE). The score on these components is an average of the answers to the underlying questions, where answers to negatively stated questions have been reversed.

The school under study

The school under study offers all types of education (pre-vocational to pre-university education). However, the experiment has been conducted in the biology classes of the second year of prevocational education. Up to the second year, the school does not make a distinction between practical and theoretical prevocational school. The tracking *within* prevocational school only takes place after the second year. The school under study is a typical, average-sized secondary school (around 1650 students in total) in a rural region with decreasing population numbers.

Findings

The results of the first experiment show that there is a statistically significant small effect of digital differentiation on the post-test score, which corresponds to a small effect of 0.18 of a standard deviation. Additional analysis shows that these effects are not driven by a specific group of students, but that all groups of students are represented in this effect. Overall, the experiment shows that differentiation slightly improves biology learning in a classroom setting with digital learning and weekly testing.

The second experiment confirms that educational outcomes are improved when teachers incorporate formative, multiple choice tests with extended, personalized feedback in their classes. The results are significantly higher than for students that received similar weekly tests with barely any feedback (only an indication of right and wrong answers at the end of the course). Again, the effect is similar across all performance groups. The analysis shows a large standardized effect size of 0.67 of a standard deviation.

The third experiment shows that simply digital testing without the extensive feedback also produces positive outcomes (when compared to withdrawing the intermediate digital tests altogether). These effects are, in

the literature, frequently explained by the increased amount of exposure to the material and/or the retrieval process of information that reactivates the memory (Roediger & Karpicke, 2006). The analysis shows a medium standardized effect size of 0.45 of a standard deviation.

Since no long-term effects (one year later) are observed for any of the three experiments, it seems that the motivation mechanism may be an important part of the explanation for the effect observed, rather than the learning productivity enhancing mechanism ('testing effect') stressed by educational scientists. Apparently, the gains were strongly tied to the specific topics discussed. Alternatively, or additionally, testing and feedback may have functioned as a way of 'teaching to the test' by the teacher of the experimental period (a different one than the subsequent year).

The importance of the motivation aspect was also highlighted in group discussions after the experiment. These revealed that treatment group students substituted some time they would otherwise spend on homework for time preparing the formative tests. The latter means that there was no radical increase in homework effort.

Cost-effectiveness

With respect to costs, the additional effort of the teacher in developing the digital formative tests is a factor that needs to be taken into account when thinking of implementation. However, it was a one-time effort (as these tests can be used again the following year). The question is whether most teachers would have to prepare these tests at all, as they are often provided by the publisher when digital learning material is purchased by a school. The cost of answer-driven feedback is probably higher, as the teacher will have to develop the algorithm in the computerized learning environment. However, this also is a one-off cost.

Although the additional effect of providing feedback on wrong (and correct) answers is definitely present, providing information via formative tests is, in itself, already very effective, at a much lower cost. Especially in this era, in which computerized education is becoming more common, additional information provision via these types of formative test is relatively easy to implement at a low cost, as soon as schools begin working with digital learning materials.

Synthesis

The eight experiments described above all provide insight into the effectiveness of using ICT in Dutch secondary education. This section provides a synthesis of all the findings from these experiments, by making comparisons and pointing to similarities and differences between the studies and the findings. This is performed from both an effectiveness point of view and with respect to the role of the teacher.

To make the comparison between the experiments easier, an overview of the previously discussed experiments is provided in Table 4.1. All experiments took place in middle school, for students from grades 7, 8 and 9. Except for the GotIT?! study, they were all randomized, although two were randomized at the class level. The experiments covered the topics of mathematics, language and biology and all took place between the school years 2012-2013 to 2014-2015. The Mousework and GotIT?! studies lasted for four months, the digital differentiation experiment for three months, and the other experiments for only six to eight weeks. The standardized effect of the experiments varies between 0.1 and 0.67. Here, it should be noted that the small effect sizes of 0.11 and 0.16 are found for language and for very short mathematics experiments. The large effect sizes of between 0.40 and 0.60 are found for long mathematics experiments, where there is a statistically significant difference between the treatment and control condition and for the testing and feedback experiments, where large effects have also been found in the literature. The overall conclusion that can be drawn from Table 4.1 reflects that found in the literature, namely that positive effects are found for mathematics, but not for language, except for spelling (see e.g. Bulman & Fairlie, 2015; Haelermans & Ghysels, 2017b). However, there are differences in the effect by age and performance level. Furthermore, positive effects for differentiation, for digital testing and for digital feedback are found, which are also similar to those found in the literature (see e.g. Roediger & Butler, 2011; Roediger & Karpicke, 2006).

As for the role of the teacher, it can be concluded as being mixed, but that this has not affected whether a statistically significant effect was found in these experiments. However, it can be concluded that the influence of the teacher is very important in the effectiveness of digital ICT tools such as those described above, as it is a crucial aspect in motivating the student to use the ICT application. In the Mousework experiments, positive effects for

Table 4.1   Overview of previously discussed experiments.

Experiment	School year	Duration	Rand-omized	Control condition	Grade	Subject	Effect	Std. effect
Mousework 1	2012-2013	4 months	Class level	Nothing	7	Math	Yes	0.40
Mousework 1	2012-2013	4 months	Class level	Nothing	7	Language	No	N/A
Mousework 2	2013-2014	4 months	Individual level	Vocabulary	7	Spelling	Yes	0.16
Mousework 3	2014-2015	4 months	Individual level	Other domains	7,8 and 9	Math	Yes	N/A
Mousework 3	2014-2015	4 months	Individual level	Other domains	7,8 and 9	Language	Yes, but only spelling and grammar	N/A
GotIT?!	2013-2014	4 months	No	Traditional teaching	7,8 and 9	Math	Yes, but only seventh grade	0.60
SMARTboard	2013-2014	6 weeks	Class level	Traditional teaching	7	Math	Yes	0.11
Digital differentiation	2012-2013	12 weeks	Individual level	No differentiation	8	Biology	Yes	0.18
Digital testing	2012-2013	8 weeks	Individual level	No testing	8	Biology	Yes	0.67
Digital testing and feedback	2012-2013	6 weeks	Individual level	Testing, no feedback	8	Biology	Yes	0.45

mathematics and spelling are found, regardless of the often negative attitude of the teacher. However, larger effects would probably have been found if teachers had stimulated students more often to practice. As regards the SMARTboard and biology experiments, the teachers had a positive attitude and knowledge level. Teachers in the SMARTboard experiments received a TPACK training before taking part in the study, and the teacher in the biology experiments developed most of the material himself and also developed the experimental study. It is possible that these effects would have been smaller if the teacher had less affiliation with ICT.

Overall, from this chapter on experiments in Dutch secondary education we can conclude that digital learning materials used in order to differentiate among students are very promising for practicing basic skills such as mathematics, language and biology, provided that the skill can easily be automated and students are stimulated to use the application to its full extent. Since students need motivation to practice or use the ICT application, other actors, such as teachers and/or parents, also play a significant role in the potential effectiveness of ICT tools.

External validity and generalization of results

The results presented in this chapter are found in the Netherlands, in some cases using specifically Dutch ICT tools, and the generalizability of these results to other ICT tools, to other educational sectors, and/or to other countries, deserves some discussion.

For generalizability, it is initially important to discuss what most likely drives the results. Is it this specific ICT (tool) in this specific setting, or is it the underlying features of the ICT (tool) that leads to its effectiveness? The studies discussed above give some indication that the effectiveness of the ICT tool or application does not come from the tool itself, but rather from specific features of the tool, such as the adaptive nature of Mousework and GotIT?!. This can also be concluded from some studies in the literature that were conducted in countries other than the Netherlands. Therefore, the Mousework and GotIT?! results, as well as the results found in the literature about similar tools, most likely also apply to other ICT tools in other countries, as long as those tools have similar characteristics as other intelligent tutoring systems. These characteristics are the didactical principles of the tools, namely: the adaptive nature of the tool, making small steps ('scaffolding'), action (practicing) and variation, direct feedback and the teacher as the supervisor.

Second, all the experiments have taken place in Dutch secondary education, in middle schools, and it is hard to argue that these results will also automatically apply to primary education and high school students. How-

ever, the literature discussed in Chapter 2 does give some indications that these types of ICT tools might also work in primary education. Furthermore, the literature also indicates that similar ICT tools can be similarly effective in other (western) countries, suggesting that these results are not limited to the Netherlands.

In regard to the comparison with Sweden, it is important to note that, although none of the studies in the literature review are from Sweden, this does not mean that ICT in education is not a topic that is directly or indirectly studied. To the best of my knowledge, there are currently one or two projects being carried out that touch upon this topic, and there might be more of which I am not aware. Some examples of the former are the projects by Hillevy Lenz Taguchi, who uses computerized attention training to study preschool children´s attention, language and communication skills[17] and the experimental studies by Torkel Klingberg also use ICT tools for children to work with, in their studies focusing on the working memory.[18]

---

17.   http://www.buv.su.se/english/research/research-projects/early-childhood-education/enhancing-preschool-children-s-attention-language-and-communication-skills-1.209094
18.   http://cognitionmatters.org/research/

Chapter 5.
Concluding Remarks

In this report, I have first summarized the knowledge from the academic literature on the effects of ICT in education. Following this, I have described the educational system in the Netherlands and the Dutch policy on ICT in education, as well as the extent to which ICT was used in 2015 and which types of ICT were used. I also made a comparison between Sweden and the Netherlands with respect to the educational system and ICT use and presence in education and student performance, the latter both based on the most recent PISA results. Lastly, I have described several small- to medium-sized experiments in which the effects of specific ICT use in education in secondary education in the Netherlands were studied.

Conclusions from the literature

The main conclusions from the international literature in economics on the effects of ICT in education are the following:

- The general investments in ICT in education without a specific purpose on what to invest in or how to use ICT in education provide mixed results, at best.
- Studies on the effects of computer-assisted instruction versus traditional classroom learning, where ICT is a complement to the teacher, do find positive effects, though very small ones.
- For specific digital learning tools, positive effects are found in developing countries, both for mathematics and for language. For Western countries, positive effects are found only for mathematics, but not for language training.
- In sum: the effectiveness of ICT in education is highly dependent on how it is used and which (pedagogical) purpose it is intended to serve.

- Cost-effectiveness of ICT (tools) in education is rarely studied, so more research is needed to draw reliable conclusions on this topic. However, the few studies that look into this conclude that the ICT (tool) is cost-effective and similar to or cheaper than reducing class size or hiring an additional teacher.
- There are many barriers to technological change for teachers, which might explain why effective technology adaptation in schools has not, to date, lived up to expectations. Teachers are either resisting the technological change in general, due to, for instance, their internal beliefs, or do not know how to apply the technology effectively in class, due to factors such as lack of time, knowledge or training.
- The closely related literature strand on change management may provide some good suggestions on how to implement changes in public sector organizations such as schools. Examples of these are providing a plan for implementation, providing resources and ensuring top-management support.

Conclusions from experiments in the Netherlands

In addition, the main conclusions from eight different randomized experiments on ICT in secondary education in the Netherlands, described in this report, are the following:

- Positive effects for mathematics and some aspects of language: The overall conclusion drawn from these experiments on ICT in education in the Netherlands is that medium positive and statistically significant effects are found for mathematics, and for some aspects of language learning. This finding on mathematics is similar to that found in the literature, but, for language, this finding differs from the literature, where no effects on language are found in developed countries.
- Individualization is effective: From the experiments, it can be concluded that it is rather the individualization of exercises that makes digital practice tools effective, and not merely the additional practice time in itself. However, students that practice more also experience more statistically significant effects.
- Effective for instructions that are easy to automate: The results indicate that digital tools that are used for adapting instructions to apply

for each student are very promising when it comes to practicing basic skills in mathematics, language and biology, such as addition and multiplication for mathematics and spelling for language. Important conditions are that the instructions are easy to automate and that students are stimulated to use the digital tool to its full extent.

- Effects differ across domains: The results indicate that students benefit differently from using digital tools depending on the age of the student and depending on the domain the tools are used in. In general, the easier domains of mathematics, spelling and grammar are predominantly beneficial for students from grades 7 and 8, and the more difficult domains of mathematics are more beneficial for higher performing students and students in grades 8 and 9.

- Effects differ across performance levels: There are statistically significant differences in the effect of adaptive digital learning materials between low-, middle-, and high-achieving student groups. Although it is found that the overall mathematics skills of both low- and middle-achieving students significantly increase when they practice with the online tool, it is significantly less effective for middle achievers than for low achievers. For high achievers, it is unclear whether there is an overall effect (although, as noted in the previous conclusion, positive effects are found for the more difficult mathematics domains). The differences in effect between low- and middle-achieving students are also seen in the separate analyses among mathematics skills domains (numbers, proportions, and measurement). Low-achieving students benefit significantly from the intervention in all mathematics domains, while middle-achieving students benefit significantly from the intervention in the domains, proportions and measurement.

- In-class-level differentiation through use of hardware is effective: Statistically significant and positive effects are found by using hardware such as interactive whiteboards in class in lower secondary education. However, these effects primarily seem due to the differentiation that was possible because of the use of the interactive whiteboard. Furthermore, a crucial part of this study was that teachers were properly trained beforehand.

- Digital tests are effective: Positive effects are found using digital tests as a learning tool, which is similar to that found in the litera-

ture, where the effects are often explained by the increased amount of exposure to the material and/or the retrieval process of information that reactivates the memory. One experiment shows that simply digital testing, even without extensive feedback, already produces positive outcomes.

- Digital feedback is effective: In addition to the previous point, even higher positive effects are found if digital feedback is given when testing digitally. Educational outcomes are improved when teachers incorporate formative, multiple-choice tests with extended, personalized feedback in their classes. The results are significantly higher than for students that were given similar weekly tests with barely any feedback. The effect is similar across all performance groups.
- Effects are often dependent on the teacher: The effectiveness of ICT in education is, in many cases, related to the way the teacher implements the ICT innovation and the knowledge of the teacher on how to use the ICT innovation, making the role of the teacher important in the effectiveness of ICT in education.
- Parental involvement is important: Parents can play an important role in stimulating students in lower secondary education to practice with adaptive online tools. A statistically significant effect is found for parental involvement on student practice behavior, and, in turn, on mathematics performance for students in grades 7 and 8. The effects of parental involvement are specifically present for low-SES students.
- To summarize: The Dutch experiments show that the effectiveness of ICT in education depends on how it is used and on the pedagogical purpose of the digital tool.

Discussion

This report shows that ICT in education can be effective, although the effectiveness of ICT in education is primarily dependent on the way ICT is implemented, how ICT is used, and on the types of learning for which ICT in education is used (the easier to automatize skills). Therefore, when ICT is used in schools, it is not only important to ensure that ICT is used at all,

and that school leaders and teachers see the need of ICT and feel confident in using it, but also that it is used in an effective way. Simply having access to ICT in education will not necessarily lead to an effective use of ICT in education. It is also important to see ICT as a tool, and not as a goal in itself.

ICT can be used in various ways, and often additional benefits of simply using ICT to increase student performance or as additional (organizational) support for the teacher are overlooked. If teachers use ICT as an adaptive practice system for their students, they can use the data (also called learning analytics) generated by the ICT system to differentiate more between students, as the learning analytics can give a good insight into the ways in which students would need extra attention and for what topics. On the other hand, the data generated by ICT use in class can also be used to monitor student performance more easily, without having to increase the number of (national) tests written by students, which can, in turn, make it easier to evaluate the teachers' role in student performance.

This report also shows that, in both countries of focus in this report, Sweden and the Netherlands, ICT seems to be very much present in schools, compared to other countries, as the PISA 2012 results show. However, educational performance is not as high as the countries would like, although this was more the case in Sweden than in the Netherlands. The question is whether this is the case in spite of the presence of ICT in education, or whether it is not used in an efficient way and therefore does not (sufficiently) contribute to educational performance. It is possible that it is not used in the most effective way, either because it is not used (even though it is present), or because it is a matter of not knowing how to do it. It also seems that perhaps a national policy is lacking, not sufficiently present or not good enough, especially because of the decentralization of educational responsibilities and choices in both Sweden and the Netherlands, but it also seems that (some) schools apparently do not feel the need or have the confidence to use ICT (at all, but most importantly, in an effective way).

Overall, increasing effective ICT use in education primarily seems a matter of: 1) increasing knowledge regarding which ICT applications in education are effective, 2) creating common support for using ICT in education and the way in which it is used, among those that actually have to use the ICT application, preferably bottom up, and 3) facilitating that schools, school managers and, most importantly, teachers, become acquainted with the ICT application and feel confident using it in the most effective way. It

is important to note that one computer per student does not seem to be necessary to achieve the positive effects found in the literature and the Dutch experiments. Change management theories and suggested approaches can play a significant role in the second and the third aspects. Furthermore, the role that other actors such as parents can play should not be underestimated, thus, in creating common support, stakeholders other than managers and teachers should not be forgotten.

Therefore, the main question is what Dutch and Swedish educational institutes can learn from this report. As regards the Netherlands, some schools are already working with adaptive digital learning tools or other types of ICT. However, for both the Netherlands and Sweden, it holds that the statistics show that most ICT in education is still not being used to teach in a different way, or being able to differentiate better, but, instead, generally for administrative reasons. It is important to emphasize that the way in which ICT is being used is the most important aspect, not merely having access to ICT. Furthermore, the experiments described in this report also show that even the Dutch schools that are working with effective ICT tools often do not use it in the most effective way, because teachers do not support the decision to work with these tools or do not know how to use the tool most effectively. Therefore, the most important aspect is creating common support and making it easier for teachers to use ICT in the most effective manner.

I believe there are three routes that can, and should, be used to try to establish this:

1. For future teachers, the route to more effective ICT use in education is via teacher-training programs, where teachers can be made familiar with different effective types of ICT that can effectively be used in education, which will most likely make them more receptive to its use and more prone to seeing the effective application of ICT as a permanent aspect of education that they can and should use.

2. For the current teaching body, the route to more effective ICT use is more via national (government) agencies and school (boards) to spread the word regarding effective ICT tools, make teachers receptive to its use, and adequately address the issue of the potential resistance by the teaching body to use ICT tools in class.

3. Furthermore, it is important to not only focus on distributing ICT in itself (infrastructure), but simultaneously to also present potential approaches for how to effectively use ICT in education (the application), and to make someone responsible for ICT use in the organization, and facilitate and support them.

As regards increasing effective ICT use by the current teaching body, the Netherlands has already taken a first step in doing this, with the foundation of the Netherlands Initiative for Educational Research (abbreviation in Dutch: NRO) in 2012. Besides funding research, among which a large part is practice based research,[19] another main responsibility of the NRO is called 'utilizing knowledge'. In other words, making sure research results reach those who should use them as a basis for their decisions. For research on ICT in education, NRO has joined forces with *Kennisnet*, an organization that aims to provide for a national ICT infrastructure in education, advises sector organizations and aims to share their knowledge with educational institutions in primary, secondary and vocational education. NRO and *Kennisnet* jointly organize an annual conference aimed at practitioners in education called 'Knowing what works with ICT in Education' (literal translation from Dutch). Another aspect in which NRO and *Kennisnet* have joined forces is by establishing a mechanism, called the 'knowledge circle' ('*Kennisrotonde*', in Dutch). Here, educational practitioners can submit questions related to effectiveness, which are answered by experts in the field, either from literature or by conducting new research.

I believe the foundation of the NRO in the Netherlands, the organization of conferences as described above, in which research results are presented to educational stakeholders, and also the example of the knowledge circle, are very good developments, although the actual realization of the knowledge utilization part has only just begun. It is of utmost importance to continue the focus on utilization of knowledge and to ensure that the results of scientific research reach the educational practitioners (managers, teachers, etc.), in such a way that those enthusiastic about ICT in educa-

---

19. The third (and largest) Mousework study, encompassing three schools, that was discussed in this report, was funded by the NRO.

tion have all the information available to make a well-considered decision regarding their investment. The next step is for these practitioners to ensure that they know how to do it and feel confident about doing so, and for the government to ensure these ICT uses are properly evaluated and the word is disseminated regarding the positive results that are found. In this light, the ICT developers and company owners also have the responsibility to, on the one hand, make their products widely available at affordable rates, and, on the other, be willing to participate in scientific research and make the data generated by their ICT system available for this end. Moreover, further research on the effective use of ICT in education is also needed, in Sweden, but to a lesser extent also in the Netherlands, as there are still many aspects of ICT use in education where sound scientific knowledge (worldwide) is scarce (such as experimental studies on learning analytics and how to effectively use those in class, and how to effectively use smartphones in class). For Swedish practice it would be useful to experimentally study the effects of the current ICT practice in education.

Providing stronger guidance on the effective use of ICT in education in Sweden and putting it on the agenda, as *Kennisnet* and NRO do in the Netherlands, is an idea that could potentially be part of the mandate of the Swedish National Agency for Education (Skolverket). Other possible actors in this, also to provide the means for more impact studies, could be the Swedish Institute for Educational Research (Skolforskningsinstitutet) and the Swedish Research Council (Vetenskapsrådet).

It could also be interesting to use other ICT-focused countries such as Singapore, China, and many other Asian countries, as examples in developing stronger national policies on how to effectively use ICT in education. Singapore, for example, has had a masterplan for ICT in Education since 1997, which roughly consists of four strategies: bringing ICT into the core of the education process, improving teachers' ICT skills, improving the sharing of best practices among educators, and upgrading schools' ICT infrastructure to keep up with developments (https://www.moe.gov.sg).

Recommendations

Based on the above, the (policy) recommendations can be organized into three categories: recommendations at the school level, recommendations at the national level and recommendations as the general level.

School level

- Deliberate choice and introduction of ICT tools: Schools should ensure, and facilitate, that information is gathered about new ICT tools and under what circumstances they are likely to be effective. This should be based on scientific research, and what form of implementation and application is needed in order to make it as effective as possible. Schools should also allow, or even make sure, that research simultaneously takes place, when new ICT tools are tried out, to learn about the effectiveness of this specific application in Sweden.
- Continuous training for professional development: The head of schools and school managers should give a more prominent role to teacher development in general and provide development opportunities for all teachers, if not present already, as this will further develop teachers' skills and mindsets in order to choose the best pedagogical methods for their teaching practices.
- Bottom-up approach: Schools should support an enthusiastic teacher with innovative ideas on how to effectively apply ICT in the classroom as a learning, instead of an administrative, tool. A well-informed and dedicated teacher will spread the word first to colleagues teaching the same subject, who will then spread the word to colleagues teaching different subjects, culminating in the involvement of nearly the entire school. From one school comes another, and so on.

National level

- More evidence about the effects of digital tools: More scientific research with causal research designs is needed in order to study the effects of different ICT uses in class, in Sweden or other Nordic countries with similar education systems, as there is currently limited evidence available regarding what works and does not work with respect to ICT in these countries.

- National knowledge system/infrastructure: There is a significant role for the national government in Sweden to not only make education practitioners enthusiastic about using ICT in the classroom, but also different ways to use it effectively, and to participate in research about this topic. The government should ensure that the knowledge is disseminated that it is not only about having ICT (devices), but also about how ICT in education can be used effectively, and which proven effective choices are available, particularly because not all ways of using ICT are effective.

- National knowledge system/infrastructure: The national government in Sweden can also learn from the Netherlands in the sense that it needs national organizations that have the aim of reaching the educational practitioners and informing them about knowledge from scientific research, as well as to stimulate research on effects of ICT in education, and/or provide stronger guidance on the effective use of ICT in education, for example by making it part of the mandate of the Swedish National Agency of Education.

- Communication of research results: The national government has a responsibility for making the results from scientific research available in accessible language to all potential stakeholders in the field of education. In other words, ensuring that research results about the effective use of ICT in education reach those who will be required to make decisions regarding its use.

- Integrating ICT in teacher education: The national government also has the opportunity to ensure that student teachers learn more about different ways of using ICT in their everyday educational practice, by making the effective use of ICT and different ways of using it in a pedagogical way part of teacher-training programs at the higher education institutes. The current teaching body should also be stimulated and facilitated to participate in these courses as part of life-long learning.

- Integrating ICT in educational plans: The national government can also stimulate that teachers at least think critically on how they can use ICT in their classes in an effective way, and make well-considered decisions regarding ICT use, by making the ways in which ICT is used part of the considerations for teachers writing their educational plans for each class.

General level

- Focus on effective practice: The head of schools, school managers and national governments can, and should, stimulate teachers to start using ICT in an effective way, while again allowing for research to take place at the same time, in order to learn about the effectiveness of this specific application in Sweden. This would entail not only focusing on using ICT, but more specifically on *how* to effectively use ICT in education, for example by looking into how ICT can help teachers accomplish their goals more effectively and efficiently.

- Do not underestimate the role of the human factor: An important – and often overlooked or underestimated – aspect of ICT in education is the human factor: teachers' and school leaders' negative beliefs and attitudes towards ICT (and perhaps towards change in general). Introducing ICT is not only about having the tools and providing the teachers with the right training about how digital tools work and can be used, it is also to a large extent about whether the teachers believe that digital tools will improve the education. If the teachers are skeptical and do not see the use of ICT in class this may be an important barrier. A successful implementation of digital tools therefore also required good leadership.

References

Abbitt, J. T. (2011). An Investigation of the Relationship between Self-efficacy Beliefs about Technology Integration and Technological Pedagogical Content Knowledge (TPACK) among Preservice Teachers. *Journal of Digital Learning in Teacher Education*, 27(4), 134–143.

Angrist, J. D., & Lavy, V. (2002). New Evidence on Classroom Computers and Pupil Learning. *Economic Journal*, 112, 735–765.

Arroyo, I., Park Woolf, B., Royer, J. M., Tai, M., & English, S. (2010). Improving Math Learning through Intelligent Tutoring and Basic Skills Training. In V. Aleven, J. Kay, & J. Mostow (Eds.), *ITS 2010, Part 1, LNCS 6094* (pp. 423–432). Berlin Heidelberg: Springer-Verlag.

Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A Review of Estimates of the Schooling/Earnings Relationship, with Test for Publication Bias. *Labour Economics* 6, 453–470.

Bando, R., Gallego, F., Gertler, P., & Fonseca, D. R. (2016). *Books or Laptops? The Cost-effectiveness of Shifting from Printed to Digital Delivery of Educational Content*. Instituto de Economica Pontificia Universidad Catolica de Chile. Documento de Trabajo 474.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics*, 122(3), 1235–1264.

Barrera-Osorio, F., & Linden, L. (2009). *The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia*. Policy Research Working Paper 4836, Impact Evaluation Series No. 29, The World Bank.

Barrow, L., Markman, L., & Rouse C.E. (2009). Technology's Edge: The Educational Benefits of Computer-aided Instruction. *American Economic Journal: Economic Policy*, 1(1), 52–74.

Bartelet, D., Ghysels, J., Groot, W., Haelermans, C., & Maassen van den Brink, H. (2016). The Differential Effect of Basic Mathematics Skills Homework via a Web-based Intelligent Tutoring System across Achievement Subgroups and Mathematics Domains: A Randomized Field Experiment. *Journal of Educational Psychology*, 108(1), 1–20.

Becker, H. J. (1990). *Effects of Computer Use on Mathematics Achievement. Findings from a Nationwide Field Experiment in Grades Five to Eight. Classes: Rationale, Study Design, and Aggregate Effect Sizes*. Baltimore, MD.

Beeby, C. E. (1966). *The Quality of Education in Developing Countries*. Harvard Univeristy Press.

Beland, L.-P., & Murphy, R. (2016). Ill Communication: Technology, Distraction & Student Performance. *Labour Economics*, 41(1), 61–76.

Bergman, P. (2015). *Parent–child Information Frictions and Human Capital Investment: Evidence from a Field Experiment*. Retrieved from http://www.columbia.edu/~psb2101/BergmanSubmission.pdf.

Beuermann, D. W., Cristia, J. P., Cruz-Aguayo, Y., Cueto, S., & Malamux, O. (2012). *Home Computers and Child Outcomes: Short-term Impacts from a Randomized Experiment in Peru*. Inter-American Development Bank Working Paper No. IDB-WP-382.

Bingimlas, K. A. (2009). Barriers to the Successful Integration of ICT in Teaching and Learning Enviroments: A Review of the Literature. *Eurasia Journal of Mathematics, Science & Technology Education*, 5(3), 235–245.

Boekaerts, M. (2002). The Online Motivation Questionnaire: A Self-report Instrument to Assess Students' Context Sensitivity. In P. R. Pintrich & M. I. Maehr (Eds.), *The Online Motivation Questionnaire: A Self-report Instrument to Assess Students' Context Sensitivity* (pp. 77-120). New York: JAI/Elsevier Science.

Borman, G. D., Benson, J. G., & Overman, L. (2008). A Randomized Field Trial of the Fast ForWord Language Computer-based Training Program. *Educational Evaluation and Policy Analysis*, 31(1), 82–106.

Bulman, G., & Fairlie, R. W. (2015). *Technology and Education: Computers, Software and the Internet*. CESifo Working Paper No. 5570.

Burns, M. K., Kanive, R., & DeGrande, M. (2012). Effect of a Computer-delivered Math Fact Intervention as a Supplemental Intervention for Math in Third and Fourth Grades. *Remedial and Special Education*, 33(3).

Cabus, S. J., Haelermans, C., & Franken, S. (2017). SMART in Mathematics? – Exploring the Effectiveness of Teaching Low Math Proficiency Students with an Interactive Whiteboard. *British Journal of Educational Technology*, 48(1), 145–161.

Cafolla, R., & Knee, R. (1995). *Factors Limiting Technology Integration in Education: The Leadership Gap*. Proceedings. Society for Information Technology & Teacher Education International Conference, 1995 Publisher: Association for the Advancement of Computing in Education (AACE), Chesapeake, VA.

Carrillo, P., Onofa, M., & Ponce, J. (2010). *Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador*. Inter-American Development Bank Working Paper No. IDB-WP-223.

Carter, S. P., Greenberg, K., & Walker, M. (2016). *The Impact of Computer Usage on Academic Performance: Evidence from a Randomized Trial at the United States Military Academy*. SEII Discussion Paper 2016.02.

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(3), 198–215.

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.

Christensen, C. A., & Gerber, M. M. (1990). Effectiveness of Computerized Drill and Practice Games in Teaching Basic Math Facts. *Exceptionality: A Special Education Journal*, 1(3), 149–165.

Christia, J. P., Czerwonko, A., & Garofalo, P. (2014). Does Technology in Schools Affect Repetition, Dropout and Enrollment? *Journal of Applied Economics*, 17(1), 89–112.

Christia, J. P., Ibarraran, P., Cueto, S., Santiago, A., & Severin, E. (2012). *Technology and Child Development: Evidence from the One Laptop per Child Program*. Inter-American Development Bank Working Paper No. IDB-WP-304.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Commissie Meijerink. (2008). *Over de drempels met taal en rekenen - Hoofdrapport van de expertgroep doorlopende leerlijnen taal en rekenen*. Enschede:

Conti, G., Heckman, J. J., & Urzua, S. (2010). The Education-health Gradient. *The American Economic Review*, 100(2), 234–246.

De Witte, K., Haelermans, C., & Rogge, N. (2015). The Effectiveness of a Computer-assisted Math Learning Program. *Journal of Computer Assisted Learning*, 31(4), 314–329.

De Witte, K. and N. Rogge (2014). Does ICT Matter for Effectiveness and Efficiency in Mathematics Education? *Computers & Education*, 75(1), 173–184.

Demetriadis, S., Barbas, A., Molohides, A., Palaigeorgiou, G., Psillos, D., Vlahavas, I., Tsoukalas, I., & Pombortsis, A. (2003). "Cultures in Negotiation": Teachers' Acceptance/Resistance Attitudes Considering the Infusion of Technology into Schools. *Computers and Education*, 41(1), 19–37.

Diamond Hicks, S. (2011). Technology in Today's Classroom: Are You a TechSavvy Teacher? *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 84(5), 188–191.

Ertmer, P. A. (1999). Addressing First- and Second-order Barriers to Change: Strategies for Technology Integration. *Educational Technology Research and Development*, 47(4), 47–61.

Falck, O., Mang, C., & Woessman, L. (2015). *Virtually No Effect? Different Types of Computer Use and the Effect of Classroom Computers on Student Achievement*. CESifo Working Paper No. 5266.

Fernandez, S. & Rainey, H.G. (2006). Managing Successful Organizational Change in the Public Sector', *Public Administration Review*, 66(2), 168–76.

Fiorini, M. (2010). The Effect of Home Computer Use on Children's Cognitive and Non-cognitive Skills. *Economics of Education Review*, 29(1), 55–72.

Freilich Hjelle, P. (2001). *Reading between the Lines: Teacher Resistance to Change*. Dissertations available from ProQuest. Paper AAI3003638. http://repository.upenn.edu/dissertations/AAI3003638.

Ghysels, J., & Haelermans, C. (2016). *New Evidence on the Effect of Computerized Individualized Practice and Instruction on Language Skills*. TIER Working Paper Series TIER WP 16/05.

Ghysels, J., Haelermans, C., & Prince, F. (2014). *The Economics of Information in Human Capital Formation – Evidence from Two Randomized Experiments with Low Stakes Tests in Secondary Education*. TIER Working Paper Series. WP 14/27.

Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A Randomized, Controlled Study of Computer-based Intervention in Middle School Struggling Readers. *Brain and Language*, 106, 83–97.

Goolsbee, A., & Guryan, J. (2006). The Impact of Internet Subsidies in Public Schools. *Review of Economics and Statistics*, 88(2), 336–347.

Goos, M., Schreier, B. M., Knipprath, H. M. E., De Fraine, B., Van Damme, J., & Trautwein, U. (2013). How Can Cross-country Differences in the Practice of Grade Retention be Explained? A Closer Look at National Educational Policy Factors. *Comparative Education Review*, 57(1), 54–84.

Groot, W., & Maassen van den Brink, H. (2010). The Effects of Education on Crime. *Applied Economics*, 42(3), 279–289.

Haelermans, C., & Ghysels, J. (2016). *The Effect of Parental Involvement on the Use of a Digital Homework Tool and on Math and Language Performance for Secondary Students – A Randomized Field Experiment.* TIER Working Paper Series TIER WP 16/01.

Haelermans, C., & Ghysels, J. (2017a). The Effect of Individualized Digital Practice on Math Skills at Home – Evidence from a Randomized Experiment on Whether and Why it Works. *Computers and Education*, 113(1), 119–134.

Haelermans, C., & Ghysels, J. (2017b). Evaluating Didactical Interventions in Primary and Secondary Education. In G. Johnes, J. Johnes, T. Agasisti, & L. López-Torres (Eds.), *Handbook of Contemporary Education Economics*: Edward Elgar Publishing.

Haelermans, C., Ghysels, J., & Prince, F. (2015a). A Dataset of Three Educational Technology Experiments on Differentiation, Formative Testing and Feedback. *British Journal of Educational Technology*, 46(5), 1102–1108.

Haelermans, C., Ghysels, J., & Prince, F. (2015b). Increasing Performance by Differentiated Teaching? Experimental Evidence of the Student Benefits of Digital Differentiation. *British Journal of Educational Technology*, 46(6), 1161–1174.

Hardy, J. V. (1999). Teacher Attitudes toward and Knowledge of Computer Technology. *Computers in the Schools*, 13(3-4), 119–136.

Hattie, J. (2009). *Visible Learning - A Synthesis of over 800 Meta-analysis Relating to Achievement*. New York, NY: Routledge.

Imberman, S. A. (2011). Achievement and Behavior in Charter Schools: Drawing a More Complete Picture. *The Review of Economics and Statistics*, 93(2), 416–435.

Inspectie van het Onderwijs. (2001). *Resultaten van vier jaar ICT-beleid in het onderwijs*. Utrecht, the Netherlands.

Keengwe, J., Onchwari, G., & Wachira, P. (2008). Computer Technology Integration and Student Learning: Barriers and Promise. *Journal of Science Education and Technology*, 17(1), 560–565.

Kennisnet (2015). *Vier in balans-monitor 2015*. Zoetermeer: Kennisnet.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C.G, Forssberg, H, and Westerberg, H. (2005). Computerized training of working memory in children with ADHD – A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(2), 177–186.

Koehler, M. J., Mishra, P., & Yahya, K. (2007). Tracing the Development of Teacher Knowledge in a Design Seminar: Integrating Content, Pedagogy, and Technology. *Computers and Education*, 49, 740–762.

Kotter, J.P. (1995). Leading Change: Why Transformation Efforts Fail. *Harvard Business Review OnPoint*. Product number 4231.

Kulik, C. L. C., & Kulik, J. A. (1991). Effectiveness of Computer-based Instruction: An Updated Analysis. *Computers in Human Behavior*, 7, 75–94.

Käser, T., Baschera, G.-M., Kohn, J., Kucian, K., Richtmann, V., Grond, U.,Gross, M. and von Aster, M. (2013). Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. *Frontiers in Technology – Developmental Psychology*, 4(489), 1–13.

Leuven, E., Lindahl, M., Oosterbeek, H., & Webbink, D. (2007). The Effect of Extra Funding for Disadvantaged Pupils on Achievement. *The Review of Economic and Statistics*, 89(4), 721–736.

Linden, L. (2008). *Complement or Substitute? The Effect of Technology on Student Achievement in India*. InfoDev Working Paper No. 17.

MacGregor, S. K., Shapiro, J. Z., & Niemiec, R. (1998). Effects of a Computer-augmented Learning Environment on Math Achievement for Students with Differing Cognitive Style. *Journal of Educational Computing Research*, 4(4), 453–465.

Machin, S., McNally, S., & Silva, O. (2007). New Technologies in Schools: Is there a Payoff? *The Economic Journal*, 117, 1145–1167.

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2015). *Using Behavioral Insights to Increase Parental Engagement: The Parents and Children together (pact) Intervention*. Cambridge MA.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-based Practices in Online Learning: A Meta-analysis and Review of Online Learning Studies*. Washington D.C.: Center for Technology in Learning.

Ministerie van Onderwijs Cultuur en Wetenschappen. (2013). *Kamerbrief over toekomstgericht funderend onderwijs*. Referentie 690041.

Mo, D., Swinnen, J., Zhang, L., Yi, H., Qu, Q., Boswell, M., & Rozelle, S. (2013). Can One-to-One Computing Narrow the Digital Divide and the Educational Gap in China? The Case of Beijing Migrant Schools. *World Development, 46*(1), 14–29.

Mo, D., Zhang, L., Luo, R., Qu, Q., Huang, W., Wang, J., . . . Rozelle, S. (2014). Integrating Computer-assisted Learning into a Regular Curriculum: Evidence from a Randomised Experiment in Rural Schools in Shaanxi. *Journal of Development Effectiveness, 6*(3), 300–323.

Muiswerk (2013). Retrieved from www.muiswerk.nl.

Mumtaz, S. (2000). Factors Affecting Teachers' Use of Information and Communications Technology: A Review of the Literature. *Journal of Information Technology for Teacher Education, 9*(3), 319–342.

Murnane, R., & Willett, J. B. (2010). *Methods Matter*: Oxford University Press.

OECD (2015). *Students, Computers and Learning: Making the Connection.*

Onderwijsraad. (2006). *Doelgericht investeren in onderwijs*. Den Haag, the Netherlands.

Pelgrum, W. J. (2001). Obstacles to the Integration of ICT in Education: Results from a Worldwide Educational Assessment. *Computers and Education, 37*(1), 161–178.

Pilli, O., & Aksu, M. (2013). The Effects of Computer-assisted Instruction on the Achievement, Attitudes and Retention of Fourth Grade Mathematics Students in North Cyprus. *Computers and Education*, 61, 62–71.

Potocki, A., Ecalle, J., & Magnan, A. (2013). Effects of Computer-assisted Comprehension Training in Less Skilled Comprehenders in Second Grade: A One-year follow-up study. *Computers and Education*, 61, 131–140.

Punie, Y., Zinnbauer, D., & Cabrera, M. (2006). *A Review of the Impact of ICT on Learning*. Brussels.

Rakes, G. C., & Casey, H. B. (2002). An Analysis of Teacher Concerns toward Instructional Technology. *International Journal of Educational Technology, 3*(1).

Roediger, H., & Butler, A. C. (2011). The Critical Role of Retrieval Practice in Long-term Retention. *Trends in Cognitive Science,* 15(1), 20–27.

Roediger, H., & Karpicke, J. D. (2006). The Power of Testing Memory - Basic Research and Implications for Educational Practice. *Perspectives on Psychological Scienc,* 1(3), 181–210.

Rouse, C. E., & Krueger, A. B. (2004). Putting Computerized Instruction to the Test: A Randomized Evaluation of a 'Scientifically-Based' Reading Program. *Economics of Education Review,* 23(4), 323–338.

Schijf, G. M., & Schijf, T. (2014). *Handleiding Online Testprogramma's.*

Skolverket (2016). *IT-användning och IT-kompetens i skolan. Skolverkets IT-uppföljning 2015.* Skolverket, Stockholm.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective Programs for Struggling Readers: A Best-evidence Synthesis. *Educational Research Review,* 6, 1–26.

Sosin, K., Blecha, B. J., Agarwal, R., Bartlett, R. L., & Daniel, J. I. (2004). Efficiency in the Use of Technology in Economic Education: Some Preliminary Results. *Recent Research in Economic Education,* 94(2), 253–258.

Teule, P. (2012, 25 July 2012). Wat kost zittenblijven nou echt? Retrieved from http://sargasso.nl/wat-kost-zittenblijven-nou-echt/.

Tomlinson, C. A. (2004). Research Evidence for Differentiation. *School Administrator,* 61(7), 30.

Tomlinson, C. A., & Kalbfleisch, M. L. (1998). Teach Me, Teach My Brain: A Call for Differentiated Classrooms. *Educational Leadership,* 56(3), 52–55.

Tournaki, N. (2004). The Differential Effects of Teaching Addition Through Strategy Instruction Versus Drill and Practice to Students with and without Learning Disabilities. *Journal of Learning Disabilities,* 36(5), 449–458.

Van Braak, J., Elen, J., Sinnaeve, I., Clarebout, G., Tondeur, J., & Evers, M. (2010). *MICTIVO, een ICT-monitor voor het Vlaamse onderwijs. Ontwikkeling en eerste resultaten.*

Van der Schaft, A. (2008). ICT oplossing voor lerarentekort? *Didactief.*

Van Vuuren, D., & Van der Wiel, K. (2015). *Zittenblijven in het primair onderwijs. Een inventarisatie van de voor- en nadelen.* CPB Policy Brief 2015/01.

Appendix: Experimental and quasi-experimental research designs

The inclusion criteria in the previous section provide one or two quasi-experimental research designs that are included in this review chapter. In regard to an explanation of these experimental and quasi-experimental research designs, Murnane and Willett (2010) provide an excellent overview of these designs applied to education. However, I will also briefly explain these designs here.

Put simply, an experimental research design randomizes students into a treatment group that is exposed to the educational technology and a control group that is not. Randomization is important to avoid selection issues, such as the motivation to use the education technology, or only higher performing students having access to the technology. Preferably, both groups are of such a size (>400) that one can assume that the students in both groups are comparable on all imaginable (but not always measurable) characteristics. If this is the case, differences in the outcome measure after exposure to the treatment can be solely attributed to the treatment.

The key issue in an experimental design is comparability between treatment and control group. This is also the key aspect of quasi-experimental designs. These designs search for a setting in which it is plausible that the treatment and control groups would also be comparable, even if there was no experimental design and there is some selection taking place. The Regression-Discontinuity (RD) design looks for an exogenously determined cut-off point that cannot be influenced by students, for example a maximum number of slots at a certain school, where the first 300 students, based on GPA ranking, get in, but number 301 does not. If the cut-off point is truly exogenous, one can assume that students just around the cut-off, in

the example the 10 students that just got in, being ranked 291 to 300 and the 10 students that just did *not* get in, ranked 301-311, are comparable, in, for example performance, motivation, etc. The difference-in-differences approach uses the feature that a whole group of students, such as educational cohorts, are exposed to an educational change, whereas the previous cohort was not. Here, students cannot influence to which cohort they belong, as in almost all schooling systems, this is determined by their date of birth. In the difference-in-differences approach, researchers look for differences *within* cohorts, between, for example, a pre-test and a post-test, or a performance increase from one school year to the other, and differences *between* cohorts, where, in general, the latter cohort was exposed to an educational change whereas the earlier cohort was not. The difference in the difference (how much more did performance increase in the group that was exposed to the change) is then attributed to the educational change. Lastly, the instrumental variable approach tries to tackle selectivity, by finding an instrument that is related to the selection issue, but not to the outcome measure. For example, this could be the selection of children with more highly educated parents into better schools. These children will probably have a high performance level, regardless of the quality of the school, and therefore an instrument is needed that relates to school choice, but not to school performance. An instrument that is often used in the literature of school selection in Europe is distance to school, as this very often determines school choice. In the instrumental variable design, researchers first estimate the likelihood of being exposed to the educational intervention, in this case a certain school, given the instrument, in this case the distance, and then estimate the effect of being exposed to the intervention, in this case attending this specific school, on the outcome measure.

To interpret the effect that is found in a study, most authors convert their finding to a so-called standardized effect size. This is an effect size expressed in standard deviations. The standard deviation is a measure of variance, related to the average. For example, say we have an average score of 50 for a test, where the effect is 5, implying that students that were exposed to the educational change score 5 points higher than control students. The smaller the range in between which all students have scored their tests, the larger this effect would be. If all students score between 40 and 6, a 5 point increase due to the intervention is a much larger effect than if the range of scores of all students is between 10 and 90. However, in both cases the average is

50 and the effect is 5. Therefore, we need the standardized effect, in order to take into account these differences in variance in scores, as well as differences in unit of measurement. The unit of measurement for student performance in the USA is, for example, GPA, ranging from F to A+, whereas in the Netherlands scores range from 1 to 10 and in Belgium from 1 to 20. Sweden has a combined system, using F to A, which can be converted to the numerical scale between 0 and 20. The standardized effect size is also used to make effects, measured in these different grading systems, comparable.

This report shows that information and communications technology (ICT) in education can be effective under certain circumstances, and that the teacher plays a significant role in this.

ICT in education generally refers to anything that involves technology in education. This includes devices such as computers, tablets, smartphones and interactive whiteboards, but also software such as educational games and digital learning tools and all educational applications that can be found on the Internet.

The report aims to contribute to the debate on which types of ICT use in education have proven to be effective. This will be discussed from the perspective of earlier research, as well as several studies about experiments at secondary schools in the Netherlands. The aim is to bring research and practice closer together, by discussing the applicability of the findings from earlier studies and the Dutch experiments.

Carla Haelermans is an Assistant Professor in Education Economics at Top Institute for Evidence Based Education Research, Maastricht University.