NFER | Educational Research

Routledge
Taylor & Francis Group

Educational Research

Special issue: Randomised controlled trials (RCTs) in education research – methodological debates, questions, challenges
Guest Editors: Ben Styles and Carole Torgerson

ISSN: 0013-1881 (Print) 1469-5847 (Online) Journal homepage: http://www.tandfonline.com/loi/rere20

Randomised controlled trials in Scandinavian educational research

Maiken Pontoppidan, Maria Keilow, Jens Dietrichson, Oddny Judith Solheim, Vibeke Opheim, Stefan Gustafson & Simon Calmar Andersen

To cite this article: Maiken Pontoppidan, Maria Keilow, Jens Dietrichson, Oddny Judith Solheim, Vibeke Opheim, Stefan Gustafson & Simon Calmar Andersen (2018) Randomised controlled trials in Scandinavian educational research, Educational Research, 60:3, 311-335, DOI: 10.1080/00131881.2018.1493351

To link to this article: https://doi.org/10.1080/00131881.2018.1493351

Published online: 23 Aug 2018.

EDUCATIONAL RESEARCH
2018, VOL. 60, NO. 3, 311–335
https://doi.org/10.1080/00131881.2018.1493351

Randomised controlled trials in Scandinavian educational research

Maiken Pontoppidan[a], Maria Keilow[a,b], Jens Dietrichson[a,g], Oddny Judith Solheim[c], Vibeke Opheim[d], Stefan Gustafson[e] and Simon Calmar Andersen[f]

[a]Department of Health, VIVE – The Danish Centre for Social Science Research, Copenhagen, Denmark; [b]Department of Education, VIVE - The Danish Centre for Social Science Research; [c]National Centre for Reading Education and Research, University of Stavanger, Stavanger, Norway; [d]NIFU – Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway; [e]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden; [f]Department of Quantitative Methods, VIVE - The Danish Centre for Social Science Research; [g]Department of Quantitative Methods, VIVE - The Danish Centre for Social Science Research

ABSTRACT

Background: The Scandinavian countries have a long history of implementing social interventions, but the interventions have not been examined using randomised controlled trials until relatively recently compared with countries like the United States and the United Kingdom.

Purpose: The purpose of this paper is to examine the history of randomised controlled trials in Scandinavian compulsory schools (grades 0–10; pupil ages 6–15). Specifically, we investigate drivers and barriers for randomised controlled trials in educational research and the differences between the three Scandinavian countries Denmark, Norway, and Sweden.

Methods: To locate relevant trials, we performed a systematic search of four bibliographic databases and a search for grey literature. Results were combined with trials located through direct contact with researchers and government officials. A trial was included if one or more interventions were randomly assigned to groups of students and carried out in a school setting with the primary aim of improving the academic performance of children aged 6–15 in grades 0–10 in Denmark, Norway, or Sweden. We included both conducted and ongoing trials. Publications that seemed relevant were screened based on full-text versions. Data extraction included information from the included studies on grade level, study period, sample size (N), project owner, funding source, and theme. In addition, we conducted two semi-structured interviews by phone or in person with central employees in funding agencies and ministries and 25 correspondences with researchers and policymakers.

Findings and conclusion: RCTs in grades 0–10 were few in all of Scandinavia until about 2011, after which there was an increase in all three countries, although at different rates. The largest number of trials has been conducted in Denmark, and the increase is more marked in Denmark and Norway compared with Sweden. International trends towards more impact evaluations and results from international

ARTICLE HISTORY

Received 13 September 2017
Accepted 22 June 2018

KEYWORDS

Randomised controlled trials; education; school; Scandinavia

CONTACT maiken pontoppidan mpo@sfi.dk

comparisons such as PISA have likely affected the development in all countries, but while many trials in Denmark and Norway are the result of policy initiatives, only one such example in Sweden was identified. We believe the lack of government initiatives to promote RCTs in Sweden is the most likely explanation for the differences across the Scandinavian countries. Funding and coordination from the government are often crucial for the implementation of RCTs and are likely more important in smaller countries such as the Scandinavian ones. Supporting institutions have now been established in all three countries, and we believe that the use of RCTs in Scandinavian educational research is likely to continue.

Introduction

The introduction of randomised controlled trials (RCTs) within the social and educational sciences in Scandinavia is relatively recent compared with the United States of America (USA), for example. Although the Scandinavian countries have a long history of implementing social interventions, often financed through the welfare state tax system and initiated by the government, the evaluation of such interventions has not, until recently, been documented using RCT designs.

In this paper, we study the development of RCTs, examining interventions aimed at improving the academic performance of children in Scandinavian compulsory schools (grades 0–10). We examine which policies and organisational efforts paved the way for the growth in educational RCTs in Denmark, Norway, and Sweden, and we discuss possible drivers or barriers for this growth and the differences across countries.

Throughout the paper, we narrowly define Scandinavia as Denmark, Norway and Sweden, in contrast with the broader group of Nordic countries that also includes Finland and Iceland. We focus on the Scandinavian countries because they are more similar than the Nordic countries.

Background

A large number of tax-funded health, education and social interventions were initiated in the Scandinavian countries through the second half of the 20th century. An economic boom in the 1950s and 1960s made the development of the welfare state possible (Kuhnle and Hort 2004). In the 1960s and 1970s, the welfare state grew to include social transfer benefits (e.g. tax-financed maternity leave and child allowances, scholarships to high-school youth and students) and universal services (e.g. free schools, university and subsidised child care; Viby Mogensen 2010; Konnerup 2011). Although the characteristics of the welfare systems in Denmark, Norway and Sweden differ, they are all based on universalism characterised by the view that all citizens have full and unconditional social citizenship rights and status (Esping-Andersen and Korpi 1987; Kuhnle and Hort 2004).

In the aftermath of the oil crisis in the late 1970s, a fiscal crisis hit the Scandinavian countries in the 1980s and 1990s (Kuhnle and Hort 2004). Increases in the number of elderly and unemployed people put pressure on the budget, and the growth in welfare benefits slowed down. Perhaps as a consequence, there was a wave of privatisation of

public services in the 1990s and a focus on new public management. With the budget constraints followed an increasing emphasis on value for money and an interest in evidence on which interventions were the most effective.

The Programme for International Student Assessment (PISA) was established by the OECD in 1997, and the first round of testing was performed in 2000. Many countries with relatively high school expenditure did not perform as highly as they may have expected. Denmark, Norway and Sweden all spend more than average in primary and lower secondary grades (OECD 2016), while their PISA performances have been varied (with some PISA scores above, some at and some below the OECD average). The PISA results placed attention on schooling and education for many national governments, including the Scandinavian governments, and started a search for knowledge about what made the highest-performing countries so successful.

School systems

The compulsory school systems in the Scandinavian countries share many key characteristics, yet there are notable differences. Primary and lower secondary schools are free of charge in all three countries, and it is mandatory to receive a minimum of nine (Sweden) or 10 (Denmark and Norway) years of schooling. Children begin school in the autumn during the year they turn six. Table 1 shows the outline of the school system in each country.

In Denmark and Sweden, the first year of school is referred to as grade 0 (or kindergarten/preschool class), whereas children start in grade 1 in Norway. Kindergarten has been offered in Denmark since 1980 and was made mandatory in 2009, thus expanding the compulsory education from nine to 10 years. In Sweden, kindergarten is not mandatory, but attendance rates are around 98% (Skolverket 2017c).

Swedish and Norwegian students graduate after grade 9 and 10 (pupil age 15), respectively. In Denmark, students are offered a 10th grade track that is voluntary for students who have fulfilled the teaching requirement and who wish to complete an extra year of schooling before moving on to upper secondary or vocational education. Around half of all students complete grade 10 (pupil age 16).

In all three countries, the majority of schools are municipal, and the operation of public schools is a municipal responsibility. The head of the school is in charge of the administrative and pedagogical management of the school and reports to the school

Table 1. Outline of the school system in Denmark, Norway and Sweden.

Educational level[a]	Denmark	Norway	Sweden
School starting age	*6 years (grade 0)*	*6 years (grade 1)*	*6 years (grade 0)*
Kindergarten or preschool class (ISCED 0)	Grade 0 'Børnehaveklasse'		Grade 0 'Förskoleklass' (optional, approx. 98% attend)
Primary education (ISCED 1)	Grade 1–6	Grade 1–7	Grade 1–6
Lower secondary education (ISCED 2)	Grade 7–9 Grade 10 (optional, approx. 50% attend)	Grade 8–10	Grade 7–9

[a] To compare educational levels across countries, we use the International Standard Classification of Education (ISCED) categories (UNESCO Institute of Statistics 2012).
Source: Authors' original, unpublished Table.

board and the municipal council. Since 1992, Sweden has had a universal voucher system governing the allocation of students to schools in primary to upper secondary grades with few restrictions on who is allowed to operate schools. Financing is still a municipal responsibility, and private schools are not allowed to charge for tuition. The Swedish system differs from the other countries: in Denmark and Norway, private schools are allowed to charge for tuition. The share of students attending privately run schools in grade 0–10 is about 10–20% in Denmark and Sweden and less than 2% in Norway (Utredningar 2013). The political management of the school system is quite similar across the Scandinavian countries, with the state issuing guidelines concerning curricula, student tests, administrative regulations, and so on, and with educational ministries and agencies being involved in the overall supervision.

Standardised testing

Sweden implemented its first form of national tests in the 1930s, mainly with the aim to aid teacher assessment and provide a norm reference for grading (Lundahl and Tveit 2014). Currently, national tests take place in grades 3, 6 and 9 (pupil ages 9, 12 and 15) in, for example, literacy/Swedish and mathematics (Skolverket 2017a). Test results from the national tests have been available for researchers since 1998 for grade 9, 2010 for grade 3 and 2012 for grade 6.

Apart from the end of school final assessments, tests have not played a large role in Denmark and Norway, historically. Only within the last decade have regular national student tests been implemented.

Denmark introduced national mandatory tests in 2009/2010 comprising, for example, literacy/Danish tests in grades 2, 4, 6 and 8, (pupil ages 8, 10, 12 and 14) and mathematics in grades 3 and 6 (pupil ages 9 and 12). It is possible to take the tests on a voluntary basis in all grades for literacy/Danish and in grades 2–7 in mathematics. In grade 9 (pupil age 15), there are also written exams in literacy/Danish and mathematics and a number of other subjects (Undervisningsministeriet 2017a). These exam results have been available for research purposes through access to the Danish national registers since 2002 and the national test results since 2010.

In Norway, there are national tests in grades 5, 8 and 9 (pupil ages 10, 13 and 14) in, for example, reading/Norwegian, English and mathematics (Utdanningsdirektoratet 2017). These tests have been conducted in their present form since 2007. In addition, students take written exams in the final year of school (grade 10 – pupil age 15) in mathematics, English and Norwegian (main or second-language form). These data can be used for research purposes and are accessible through administrative registers.

The stated purpose of the current tests includes – for all three countries – that the tests should support teachers in their assessment of students, and thereby improve school quality and student learning. Accountability purposes – in terms of meeting targets for or following the development of student achievement – are emphasised to different degrees. In Denmark, accountability is mentioned only in relation to the national level. In both Norway and Sweden, one of the purposes of the tests is to follow the development of student achievement on the municipal and school level as well (Skolverket 2018; Undervisningsministeriet 2017b; Utdanningsdirektoratet 2017).

Educational research

Educational research in Scandinavia has, historically, been characterised by a very strong emphasis on so-called 'qualitative' research (Kunnskapsdepartementet 2006; Sundell and Stensson 2010; Broady et al. 2011; Holm and Jæger 2011; Norges 2013). The educational research environment in the three countries is closely connected to the education and further training of teachers, and a number of educational researchers have had previous careers in teaching. Teacher training in Sweden is conducted at regular universities and university colleges, whereas teacher training institutions in Denmark and Norway used to be separate institutions with a relatively small share of employees with research training. During the 1990s and 2000s, Danish and Norwegian teacher training institutions were gradually changed or integrated into university colleges and universities.

The communication with and culture among teachers are therefore likely to have influenced the educational research environment. The emphasis on qualitative research fits well with a prevailing view of teaching, where teachers, to a large degree, have been free to choose their own methods. The argument for this view – coined 'let the 1000 flowers bloom' (Deding and Høg 2015) – is that every individual is unique and that the individual teachers or schools must be allowed to decide freely which method or intervention they find most suitable in each case. This local anchoring of methods may have been difficult to combine with methodological transparency and rigour (Deding and Høg 2015), particularly in impact evaluations.

Institutions aiming to support and disseminate educational research, akin to the Institute of Education Sciences in the USA, have been established in all three countries during the last 15 years. The Nordic Campbell Centre was established in Denmark in 2002 by a grant from the Ministry of Social Affairs, and The Danish Clearinghouse for Educational Research was established in 2006 at Aarhus University. Both centres produce reviews of educational research. The main office of the international Campbell Collaboration moved to Norway in 2007 and is funded by the Norwegian Institute of Public Health, although the Campbell Collaboration produces systematic reviews within areas such as education, crime and justice, and social welfare.

TrygFonden's Centre for Child Research was established in 2013 at Aarhus University, Denmark, and was funded by the nonprofit foundation TrygFonden. The centre carries out interdisciplinary research projects and specialises in scientific impact evaluations in the fields of social policy, education, and crime prevention, focussing on children and young adults.

The Knowledge Centre for Education was established as a department of the Research Council of Norway, Division for Society and Health in 2013. The primary goal of the Knowledge Centre is to produce, gather, synthesise and disseminate knowledge from research on issues relevant to the education sector in order to contribute to a knowledge-based policy development, management and practice, and to enlighten and provide a knowledge base for the educational debate.

In Sweden, the Institute for Evaluation of Labour Market and Education Policy was recently given the formal responsibility of evaluating the effects of education policies (Statens Offentliga Utredningar 2013), and it also gives grants to educational research. The Institute for Educational Research was established in 2015 and aims to validate educational research in terms of quality and relevance, perform systematic reviews, disseminate research results for educational practitioners, identify areas where relevant

research is needed, and provide funding for research projects (Skolforskningsinstitutet 2017).

In sum, during the recent decades, teacher training institutions in the Scandinavian countries were integrated into university colleges and universities, several institutions funding, conducting, and disseminating educational research were established, and standardised testing was implemented in all three countries. This created a larger focus on educational research, allowed for larger research groups and more variation in the research methods employed, and laid the foundation for educational RCTs in Scandinavia.

Study aim

The purpose of this paper is to examine the development of RCTs of interventions aimed at improving the academic performance of children in Scandinavian compulsory schools (grades 0–10: pupil ages 6–15), to identify policies and organisational efforts that paved the way for the growth in educational RCTs in Denmark, Norway and Sweden, and look into possible drivers or barriers for this growth and the differences across countries.

Method

To collect relevant information, we performed a systematic search of the research literature, searched governmental documents, and contacted central employees in funding agencies, ministries and research environments.

Search strategy

We searched for education trials in several ways: (1) a systematic search of bibliographic databases; (2) a search for grey literature and (3) contacting researchers and government officials. We included both trials with published results and trials that were ongoing.

Databases were searched in July 2017. We searched four international bibliographic databases for publications from 1997 and onwards: PsycInfo, ERIC, Soc Index and Econ Lit. The search was made up of conjunctions of the following terms: student*, pupil*, child*, Denmark*, Sweden*, Norway*, Scandinavia*, learn*, teach*, instruction*, school*, mathematic*, education*, practice*, classroom*, RCT*, random* control* trial*, trial* and experiment*. Publications were screened based on abstract and title.

We used snowballing techniques to find grey literature (i.e. unpublished and ongoing studies). We searched Google and the homepages of ministries, government agencies, and funding bodies in the three countries. We looked through reference lists and contacted researchers who had conducted trials and asked them whether they had ongoing studies and whether they knew of other RCTs. We also contacted government and private officials involved in funding educational research.

Data collection

In order to gain knowledge about the historical reasons for funding decisions of education trials and methodological preferences among policymakers, we contacted central

employees in funding agencies, ministries, and research environments. Employees were selected for participation if they worked in a central position with knowledge of processes leading to funding decisions. We tried to get an interview with one government representative in each country. We conducted two semi-structured interviews – one in person (Denmark) and one by telephone (Norway). Both interviews lasted about 45 min and followed the same interview guide. We were unable to arrange an interview with a Swedish representative. We also gathered information through 25 email correspondences with employees from all three countries. E-mail correspondences mostly addressed specific questions about trials the employee was involved in, questions about historical development, or comments on specific text sections.

To supplement the information gathered through personal contact, we studied central documents and policies from ministries and government agencies. To learn more about the absence of policy-initiated RCTs in Sweden, we performed a search of the publication database of the National Agency for Education (Skolverket 2017b) – the government agency mainly responsible for evaluations in grades 0–9 during the period. We used the search terms 'evaluation' (utvärdering), 'intervention' (intervention) and 'effect' (effekt) for the publication database. We also searched the agency's website for mentions of RCTs using abbreviated terms synonymous with randomisation in Swedish (random* and slump*).

Eligibility criteria and data extraction

We used the following inclusion criteria for each trial: the intervention had to be randomly assigned. The trial had to examine an intervention carried out in a school setting with the primary aim of improving the academic performance of children in grades 0–10 in Denmark, Norway or Sweden. The intervention did not have to consist of solely academic activities; we included, for instance, a physical training trial with the main purpose of improving students' mathematics skills. But trials examining non-academic interventions such as anti-bullying, mindfulness and providing school lunches were excluded if they did not include academic outcomes but aimed to improve, for example, student well-being.

Each publication was screened by a research assistant who was under supervision by the first, second and third authors. Publications that seemed relevant were screened using the full-text version. Screening was performed in Mendeley. We extracted information about grade level and student age, study period, sample size (*N*), project owner, funding source and the overall aim of the included trials.

Trials were characterised as either policy- or researcher initiated. Policy-initiated trials were defined as trials initiated by and partly or fully financed by national or local government institutions. These trials are typically initiated by contracting authorities such as ministerial institutions or agencies who issue calls for tenders. Following public procurement rules, (invited) research institutions and/or private organisations will submit their tenders and compete for the contract. The level of detail and predefined requirements regarding research designs described in the calls for tenders may vary substantially between calls, i.e. some projects are predefined as randomised trials, while others are open for different designs. Researcher-initiated trials were defined as trials initiated by research institutions such as universities, research centres or university

colleges. These projects may be funded fully by the institution itself or may be partly or fully funded by, for example, research councils, foundations or ministries.

Ethical considerations

It was not necessary to obtain ethics approval for this study. We informed all participants that we sought information on educational RCTs and we explained that the information they provided would be presented in a condensed form. Interview participants were asked whether they wanted to be anonymous and whether any of the information given was confidential.

Findings

The bibliographic database search yielded 1336 articles, of which 33 were screened in full text. When combining all searches, 55 studies met the inclusion criteria. Table 2 provides information about the characteristics of each included trial. The table is divided by the country in which the trial is conducted and whether it is policy- or researcher initiated.

Figure 1 shows the development of RCTs in the three countries. For a few trials, we lack information on the timing of the project; these trials are included in the most likely time period by using publication dates.

There were few RCTs up until 2010, after which the number of initiated studies increased in all countries, particularly in Denmark. We describe the characteristics of included RCTs and document the methodological preferences among policymakers by country below.

Denmark

Number of trials

In Denmark, the first RCT was initiated in 1992 (Elbro and Petersen 2004). The trials initiated in the 1990s were smaller scale studies ($N < 100$) compared with later trials, and the majority of the early trials were initiated by researchers within the field of linguistics. In total, the search identified 14 researcher-initiated RCTs.

We identified 13 policy-initiated Danish trials. The first one was initiated in 2012 and was an evaluation of a co-teacher intervention performed for the Ministry of Education by TrygFonden's Centre for Child Research (Andersen et al. 2016). The Centre has been responsible for several of the Danish education RCTs since then.

Funding

Almost all policy-initiated studies have been fully or partly funded by the Ministry of Education (12), whereas only one trial was funded by the Danish Ministry of Social Affairs. Most policy-initiated trials were carried out at TrygFonden's Centre for Child Research (8), a smaller number at SFI – The Danish National Centre for Social Research (4), and one trial at Aarhus University.

The majority of the researcher-initiated trials were funded fully or partly by the Danish Research Councils (6). TrygFonden also funded some (4).

Table 2. Policy and research initiated RCTs in the educational sector, by country.

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Denmark, policy-initiated RCTs (*N* = 13)							
Grade 6 Age 12	The effect of co-teacher interventions (Effekten af 2-lærer)	2012–2014	10,198 students (241 schools)	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of co-teacher interventions	(Andersen et al. 2016)
Grade 4 Age 10	Effects of first language based education in classes – first trial (Modersmålsbaseret undervisning – forsøg for hele klasser)	2013–2016	1931 students (90 schools)	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of first language based education implemented in classes	(Andersen, Humlum, and Nandrup 2016)
Grade 4 Age 10	Effects of first language based education in small groups (Modersmålsbaseret undervisning – forsøg for grupper)	2013–2016	648 students (51 schools)	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of first language based education implemented in small groups	PI: S.C. Andersen and M. Humlum
Grade 4–5 Age 10–11	Effects of first language based education in classes – second trial (Modersmålsbaseret undervisning – andet forsøg for hele klasser)	2013–2016	2546 students (118 schools)	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of first language based education implemented in classes	PI: S.C. Andersen and M. Humlum
Grade 1 Age 7	Effects of first language based education in small groups (Modersmålsbaseret undervisning – forsøg med modersmåls-undervisning)	2013–2016	451 students (105 groups)	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of first language education implemented in small groups	PI: S.C. Andersen, M. Humlum and T.S. Guul
Grade 1–7 Ages 7–14	Educational support interventions for children in foster care (Skolestøtte til børn i familiepleje)	2013–2018	153 students (136 schools)	SFI – The Danish National Centre for Social Research	Danish Ministry of Social Affairs	Effect of educational support interventions for children in foster care on academic performance, cognitive development, behaviour and well-being	PI: M. Eiberg
Grade 8 Age 14	Effects of student involvement (Elevinddragelse i folkeskolen)	2014–2015	6161 students (158 schools)	Aarhus University	Danish Ministry of Education	Effects of student involvement	(Jakobsen et al. 2016)
Grade 5 and 7 Age 11 and 13	Effects of a student coping course. The Inclusion Panel – a longitudinal panel study (Inklusionspanelet. Elevrettet mestringsforløb; MinMestring)	2014–2016	10,127 students (169 schools)	SFI – The Danish National Centre for Social Research	Danish Ministry of Education	Effect of a student targeted intervention to support the inclusion of students with special needs in general education classrooms.	(Keilow et al. 2016)
Grade 5 and 7 Age 11 and 13	Effects of teacher-targeted training. The Inclusion Panel – a longitudinal panel study (Inklusionspanelet. Lærerrettet kursus i inklusion)	2014–2016	2066 students (169 schools)	SFI – The Danish National Centre for Social Research	Danish Ministry of Education	Effect of a teacher targeted intervention to support the inclusion of students with special needs in general education classrooms.	(Keilow et al. 2016)
Grade 3 and 5 Age 9 and 11	Effects of school directed interventions for children with low socioeconomic status (Skolerettede indsatser)	2015–2018	Planned sample size: 600 classes	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of school directed interventions for children with low socioeconomic status	PI: M. Rosholm
Grade 3 and 5 Age 9 and 11	Effects of Club Letterbox and Paired Reading in a Danish context (Skole-understøttende indsater for udsatte børn)	2015–2018	Planned sample size: 650 students	TrygFonden's Centre for Child Research	Danish Ministry of Education, co-funded by TrygFonden	Effects of Club Letterbox and Paired Reading in a Danish context	PI: M. Rosholm
Grade 8 Age 14	Effects of intensive learning camps (Turboforløb)	2016–2019	Planned sample size: 1800 students	TrygFonden's Centre for Child Research	Danish Ministry of Education	Effects of intensive learning camps	PI: M. Rosholm and K. Gumede

(*Continued*)

Table 2. (Continued).

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Grades 4–5 and 7–8 Age 10–11 and 13–14	Quality in Danish language and math teaching (KiDM – Kvalitet i dansk og matematik)	2016–2018	2400 students (12 schools) and 2800 students (14 schools), respectively for Danish and Mathematics interventions	University of Southern Denmark, Aalborg University, University College South Denmark, University College Zealand, University College of Northern Denmark	Danish Ministry of Education, Headmasters' union (Skolelederforeningen), Teachers' union (Danmarks Lærerforening)	Effects of a didactic method and connected teaching materials, which promotes dialogical, creative and application-oriented elements of teaching	PI: T. Illum Hansen
Denmark, researcher-initiated RCTs (N = 14)							
Kindergarten (Grade 0) Follow up in grades 2, 3 and 7 Age 4–6, 8, 9 and 13	Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia	1992–1997	123 students in total (35 treatment classes, 44 control classes), 82 at-risk pupils randomised to treatment or control	Department of General and Applied Linguistics, University of Copenhagen	Danish Research Council and Rebekka Foundation	Effect of phoneme awareness and letter sound training	(Elbro and Petersen 2004)
Grade 5 Age 11	Evaluation of synthetic speech feedback	1994–1996	65 students (4 schools)	Centre for Reading Research, University of Copenhagen	Danish Research Council and University of Copenhagen	Effects of synthetic speech feedback for disabled readers with language disorders	(Elbro, Rasmussen, and Spelling 1996)
Grade 4–5 Age 10–11	Effects of morphological awareness training on the reading and spelling skills of children with dyslexia	1994–1996	60 students	Centre for Reading Research, University of Copenhagen	Danish Research Council for the Humanities and Danish Ministry of Education	Effects of morphological awareness training on the reading and spelling skills of children with dyslexia	(Elbro and Arnbak 1996; Arnbak and Elbro 2000)
Grade 2–8 Age 8–14	Synthetic speech feedback for teaching reading to disabled readers with language disorders	1994–1996	74 students (4 schools)	Department of General and Applied Linguistics, University of Copenhagen	No external funding, University of Copenhagen	Effects of using synthetic speech feedback when teaching disabled readers with language disorders to read	(Elbro, Rasmussen, and Spelling 1996)
Grade 8 Age 14	The importance of genre knowledge for text comprehension – a training study	2002–2003	326 students (17 whole classes)	Centre for Reading Research, University of Copenhagen	Danish Research Council for the Humanities	Effect of genre knowledge training on student text comprehension	(Elbro and Knudsen 2010)
Grades 4–6 Age 10–12	Effects of it support (It-støtte i almen-undervisningen)	2009–2012	490 students	SFI – The Danish National Centre for Social Research and Horsens Municipality	SFI – The Danish National Centre for Social Research and Horsens Municipality	Effect of a three year it support intervention (CD-ord) on student reading abilities (using Danish language national test scores)	(Christensen et al. 2015)
Grades 1–2 Age 7–8	Effects of teacher targeted Classroom Management training on student concentration, well-being and learning (Efteruddannelse af lærere i indskolingen ELI)	2010–2012	1160 students (22 schools)	Centre for Strategic Education Research (CSER), University of Aarhus	Danish Council for Strategic Research	Effects of teacher targeted Classroom Management training on student well-being, concentration skills and learning	(Keilow et al. 2015; Keilow et al. 2017)
Grade 1 Age 7	A 9-month classroom-based physical activity programme involving integration of physical activity into the maths-lessons delivered by the schools' maths teachers	2012–2013	505 students	Centre of Research in Childhood Health, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark	IMK foundation, Copenhagen and University of Southern Denmark, Odense	Effect of a classroom-based PA intervention on mathematical achievement, creativity, executive function, BMI and aerobic fitness	(Have et al. 2016)
Grade 2–3 Age 8–9	Effect of a home-based reading intervention (READ)	2013–2016	2140 students (29 schools)	TrygFonden's Centre for Child Research	TrygFonden's Centre for Child Research	Effect of a home-based reading intervention	(Andersen and Nielsen 2016)

(*Continued*)

Table 2. (Continued).

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Grade 6 and 7 Age 12–13	LCoMotion (Learning, Cognition and Motion) – a school-based physical activity intervention on cognitive performance	2013–2014	632 students (16 schools)	Department of Sport Science and Clinical Biomechanics, Research Unit for Exercise Epidemiology, Centre of Research in Childhood Health, University of Southern Denmark	Danish Ministry of Education	Effects of a school-based physical activity intervention in enhancing cognitive performance in 12–14-year-old adolescents	(Tarp et al. 2016)
Grade 2 Age 8	Scaffolding 2. graders' reading of unfamiliar text with digital learning material that supports and strengthens students' decoding while students are reading for meaning	2015	1013 students (47 classes)	University College Lillebælt (Center for Anvendt forskning i Pædagogik og Samfund)	No external funding, PhD Grant from University College Lillebælt	Effects of a three month reading intervention using digital learning material	PI: S.T. Gissel
Kindergarten (Grade 0) Age 6	Integrated teaching of spelling and reading for at-risk children	2015–2016	74 students (17 classes)	Centre for Reading Research, University of Copenhagen	TrygFonden	Effect of integrated teaching of spelling and reading for at-risk children	PI: C. Elbro
Grade 2 Age 8	Digital learning material (Digitalt læringsredskab)	2016–2017	1013 students (47 classes)	TrygFonden's Centre for Child Research	TrygFonden	Effects of digital learning material	(Gissel and Andersen 2017)
Grade 2 Age 8	READ Implementation (READ Implementering)	2016–2018	Approx. 7800 students (572 schools)	TrygFonden's Centre for Child Research	TrygFonden	Effect of READ (a home-based reading intervention) implemented in several municipalities in Denmark	PI: S.C. Andersen and U. Hvidman
Norway, policy-initiated RCTs ($N = 4$)							
Grade 5–7 Age 10–12	Improving Interprofessional Collaboration in Norwegian Primary Schools. A cluster-randomised effect evaluation of the LOG-model	2016–2019	1776 students (37 schools)	Oslo and Akershus University College of Applied Sciences (HiOA) and University of Tromsø	Norwegian Directorate for Education and Training	Effect of the LOG-model: management, organisation and implementation of a development process where the goal is an appropriate multidisciplinary collaboration in the school	PI: Ira Malmberg-Heimonen
Grades 5–7 Age 10–12	Increased healthcare resource in systematic and structured interaction with school	2016–2019	8966 students (108 schools)	NIFU – Nordic Institute for Studies in Innovation and Centre for Learning Environment at the University of Stavanger and the Norwegian Institute of Public Health (FHI)	Norwegian Directorate for Education and Training	Effect of increased healthcare resource in lower secondary school on drop-out and completion in upper secondary education	PI: Roger Andre Federici
Grades 1–4 Age 6–9	Small group Instruction in Mathematics for Pupils Level 1–4: Effects of a randomised controlled trial intervention study	2016–2021	(160 schools) (Ongoing recruitment)	NIFU – Nordic Institute for Studies in Innovation, Research and Education	Norwegian Research Council, Educational Research (LÆREEFFEKT)	Effect of increased teacher-student ratio in Mathematics instruction (small group instruction) in grades 1–4	PI: V. Opheim and K.V. Salvanes
Grades 1–4 Age 6–9	Two Teachers: Increasing the Opportunities to Differentiate Literacy instruction	2016–2021	5700 students (150 schools)	Norwegian Centre for Reading Education and Reading Research, University of Stavanger	Norwegian Research Council, Educational Research (LÆREEFFEKT)	Effect of increased teacher-student ratio (co-teaching) and professional development for teachers, in literacy instruction	(Solheim, Rege, and Erin 2017)
Norway, researcher-initiated RCTs ($N = 7$)							

(*Continued*)

Table 2. (Continued).

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Grade 6 Age 11	Activation of background knowledge for inference making	2009–2012	236 students (6 schools)	University of Copenhagen and University of Stavanger	University of Stavanger	Activation of background knowledge for inference making: Effects on reading comprehension. Scientific Studies of Reading	(Buch-Iversen 2010; Elbro and Buch-Iversen 2013)
Grade 5 Age 10	Active Smarter Kids (ASK)	2013–2017	1145 students (57 schools)	Sogn and Fjordane University College	Norwegian Research Council – Strategic Projects – University Colleges (SHP), Gjensidige foundation, Norwegian School of Sport Sciences	Develop and test the effect of an intervention that stimulates increased physical activity in school	(Resaland et al. 2015; Resaland et al. 2016)
Grade 1 Age 6	On Track – Reducing the number of children with reading difficulties	2014–2018	1170 students (19 schools)	Norwegian Centre for Reading Education and Reading Research, University of Stavanger	Norwegian Research Council, Educational Research (FINNUT)	Early intensive reading intervention for children at risk of reading difficulties	(Lundetræ et al. 2017; Solheim et al. 2018)
Grade 1 Age 6	The Down's syndrome LanguagePlus-project	2014–2018	*Not reported*	Department of Special Educational Needs, University of Oslo	Norwegian Research Council, Educational Research (FINNUT)	Develop and test language intervention for first graders who have Down's syndrome	(Næss et al. 2017)
Grade 2 Age 7	The vocabulary learning challenge – How can we improve learning efficiency?	2015–2019	724 students (12 schools)	Department of Special Educational Needs, University of Oslo	Norwegian Research Council, (FRIHUMSAM)	Develop and test effect of vocabulary intervention	PI: B.E. Hagtvedt
Grade 1 Age 6	Ameliorating arithmetic problems in children that struggle with early number sense: a randomised controlled trial	2016–2019	120 students (9 schools)	Department of Special Educational Needs, University of Oslo	University of Oslo	Develop and test intervention in early maths skills	PI: M. Melby-Lervåg, P. Aunio & R. Mononen
Grade 1 and 2 Age 6 and 7	Enhancing second-language learning in young language minority students: an intervention study in early elementary school years	2016–2020	137 students	Department of Education, University of Oslo	University of Oslo	Develop and test language intervention in young minority students	PI: M. C. Heller
Sweden, policy-initiated RCTs ($N = 1$)							
Grade 0–12 Age 6–18	Interventions to improve the quality of education for newly-arrived immigrant students (Insatser för att stärka utbildningens kvalitet för nyanlända elever)	2016–2020	100 municipalities (Students and schools not reported)	National Agency for Education and Institute for Evaluation of Labour Market and Education Policy	National Agency for Education	Support for municipalities to improve the quality of education for newly-arrived immigrant students	PI: Anna Sjögren & Björn Öckert
Sweden, researcher-initiated RCTs ($N = 16$)							
Grade 4 Age 10	Why Do Some Resist Phonological Intervention? A Swedish Longitudinal Study of Poor Readers in Grade 4	1994–1995	65 students (14 schools)	Department of Education and Psychology, Linköping University	Municipality of Norrköping	Examines phonological awareness instruction given to struggling readers	(Gustafson, Samuelsson, and Jerker 2000)
Grade 2–3 Age 8–9	Phonological or Orthographic Training for Children with Phonological or Orthographic Decoding Deficits	2003–2004	100 students (20 schools)	Department of Behavioural Sciences and Swedish Institute for Disability Research, Linköping University	Bank of Sweden Tercentenary Foundation	Compares phonological training to orthographical training to ordinary special instruction for children with decoding difficulties	(Gustafson, Ferreira, and Rönnberg 2007)
Grade 3 Age 9	Effects of a randomised reading intervention study	2007–2010	112 students (59 schools)	Department of Education, University of Gothenburg	Swedish Research Council and Swedish Tercentenary Bank Foundation	Evaluates a reading and fluency training programme based on phonemic awareness for students with word decoding deficits	(Wolff 2011; Wolff 2016)

(*Continued*)

Table 2. (Continued).

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Grade 2 Age 8	Effects of Three Interventions on the Reading Skills of Children With Reading Disabilities in Grade 2	2008–2009	100 students	Department of Behavioural Sciences and Learning, Linköping University	Swedish Council for Working Life and Social Research	Effects of three computerised training programmes for children with reading disabilities: One bottom-up intervention aimed at improving word decoding skills and phonological abilities, one intervention focused on top-down processing on the word and sentence levels, and one was a combination of these two training programmes	(Gustafson et al. 2011; Fälth et al. 2013)
Grade 1–6 Age 7–12	Auditory Processing in Developmental Dyslexia: An Exploratory Study of an Auditory and Visual Matching Training Programme with Swedish Children with Developmental Dyslexia	2006–2007	41 students (3 schools)	Department of Applied Sciences of Education, University of Helsinki	Finnish Cultural Foundation and University of Helsinki, Department of Applied Sciences of Education, Special Education	Examine if training using a nonverbal auditory-visual matching task had a remedial effect on reading skills in developmental dyslexia	(Törmänen and Takala 2009)
Kindergarten (grade 0) Age 6	Numbers, reasoning, and representation – An intervention study in kindergarten mathematics	2010–2013	124 students (9 schools)	Centre for Education Science and Teacher Research, University of Gothenburg	*Not reported*	Develop and evaluate an intervention with structured and explicit instruction focussing on numbers and children's and teachers' collective reasoning about representations	(Sterner, Wolff, and Helenius 2015)
Kindergarten (Grade 0) Age 6	Phonological awareness training with articulation promotes early reading development	2013–2014	69 students (2 schools)	Linnaeus University	No external funding	Examine a whole-class intervention of phonological training with articulation	(Fälth, Gustafson, and Svensson 2017)
Grade 2 Age 8	Short and Long-Term Effects of a Mathematics Tablet Intervention for Low Performing Second Graders	2013–2016 (project started in 2011)	283 students	Department of Psychology, Uppsala Universitet	National Board of Health and Welfare (Socialstyrelsen)	Examine a tablet intervention targeting mathematics skills and working memory for low performing second graders	PI: M.H. Hallstedt
Grade 2–3 and grade 7–9 Age 8–9 and 13–15	Intensive Reading with Reading Lists – An Intervention Study	*Not reported*	60 students	Linnaeus University	No external funding	Examine an intervention where students with decoding difficulties get to train with reading lists by the 'Wendick model of intensive reading'	(Fälth, Nilvius, and Anvegård 2015)
Grade 1 Age 7	Working Memory Training – A Cogmed Intervention	*Not reported*	32 students	Linnaeus University	No external funding	Evaluate an intervention using a computerised programme for working memory training	(Fälth, Jaensson, and Johansson 2016)

(*Continued*)

Table 2. (Continued).

Educational level and average age	Project title	Study period	N	Project owner	Funding source	Theme	Publications, working paper or Principal Investigator (PI)
Kindergarten (Grade 0) Age 6	Behavior and neuroimaging at baseline predict individual response to combined mathematical and working memory training in children	Spring 2015	308 students (2 schools)	Department of Neuroscience, Karolinska Institutet	Marcus and Amalia Wallenberg Foundation	Four treatments targeting a general population of 6-year-olds: working memory training (WMT), number line training (NLT), combined WMT and NLT, and reading (active control). The three first interventions aim to improve mathematics abilities	(Nemmi et al. 2016)
Grade 2 Age 8	Traditional algorithm or decomposition method? An intervention study focusing on students' declarative, procedural, and conceptual knowledge of arithmetic	2016	390 students	Linköping University	Municipalities of Linköping and Norrköping	Examine an intervention focusing on students' declarative, procedural, and conceptual knowledge of arithmetic	PI: M. Engvall
Grade 3 Age 9	*Not reported*	2017-	*Not reported*	Department of Education and Special Education, University of Gothenburg	*Not reported*	Vocabulary intervention in third grade	PI: U. Wolff
Kindergarten (Grade 0) and grade 1 Age 6 and 7	Multi-sensory literacy learning for at-risk students in kindergarten and first grade	2017–2018	Planned sample size: 161 students (12 schools)	VIVE – The Danish Centre for Social Science Research	Institute for Evaluation of Labour Market and Education Policy	Examine the effects of a multisensory literacy intervention for at-risk students on tests of decoding skills, letter knowledge, phonological awareness, and motivation	PI: J. Dietrichson
Grade 5 Age 11	Inclusion through group learning: A mixed-methods study (Inkludering genom lärande i grupp)	2017–2018	Planned sample size: 1200 students	Department of Education, Uppsala University	Swedish Research Council	Examine an intervention where teachers get training in how to implement a co-operative learning procedure in their classrooms	PI: N. Klang
Kindergarten (grade 0) to grade 3 Age 6–9	Effective interaction for optimal language development and classroom learning (Effektiv interaktion för optimal språk-utveckling och lärande i klassrummet. Lärare och barn i en randomiserad kontrollerad studie)	2017–2019	*Not reported*	Department of Clinical Sciences, Lund University	Institute for Educational Research	Communication and literacy intervention in the early years (K-3)	PI: B. Sahlén

The number of students, schools and classes may be calculated differently in the trials and may therefore not be directly comparable.

Average student age was estimated from grade levels if not stated in the study documentation. Other types of information that we were unable to acquire at the time of this study (grade level, study period, number of observations or funding source) is indicated as *Not reported*.

PI: Principal Investigator.

Source: Authors' original, unpublished Table.

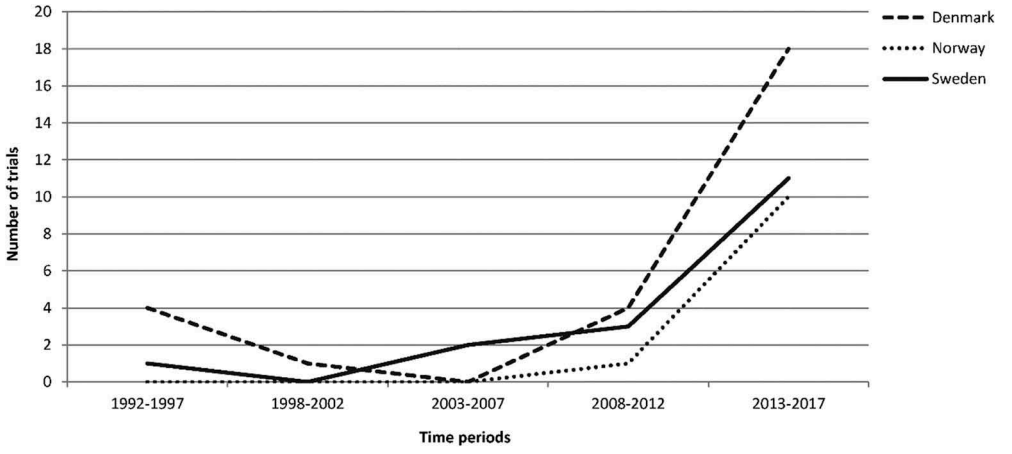

Figure 1. The development of randomised controlled trials in the three countries.
Source: Authors original, unpublished Figure.

Topics

The early educational RCTs in Denmark focussed on a narrow target group of children with dyslexia or other specific learning disabilities. As these types of studies continued to be carried out, other trials with broader scopes and larger sample sizes ($N > 1000$) were initiated alongside. These trials from 2010 and onwards include impact evaluations of more general teaching methods, such as classroom management or co-teacher interventions. Other examples include interventions to improve reading for at-risk groups, and interventions involving physical activity or digital learning materials.

Age groups

Several of the researcher-initiated trials evaluate interventions at rather early-grade levels (7 of the 14 trials cover grades 0–3, pupil age 6–9), whereas the policy-initiated trials cover a broader age range. In total, studies cover grades 0–8 (pupil age 6–14).

Methodological preferences among policymakers

In 2006, a report from the Danish Globalisation Council ('Globaliseringsrådet') recommended more quantitative education research. This recommendation was included in the Research 2015 ('Forsk 2015') goals that were published in 2008. Education research was highly prioritised by the Danish Council for Strategic Research, which reserved a significant amount of funding resources for impact studies for the years 2009 and 2010. This included funding for a trial evaluating teacher-targeted classroom management training in 2010, which was the first large RCT in Danish education research that had a different focus in comparison with previous interventions focussed specifically on reading.

The Danish School Council (Skolerådet) was established in 2006 with the aim of advising the minister of education. From 2009, the council recommended that the minister change the practice of supporting many small projects without proper impact evaluation designs to prioritise the support of larger projects incorporating impact evaluations. The School Council developed guidelines for research and development

projects in 2011. With the election in 2011, the government changed from right wing to left wing. The push for more experimental trials gained support and the guidelines from the School Council led to the development of a strategy for research and development. This was published by the Ministry of Education in 2014 and is still in use. The aim of the strategy was to support the development of high-quality knowledge on what works for schools and educational institutions to improve students' academic skills and well-being (Undervisningsministeriet 2014, 4). The strategy highlighted that the Ministry of Education would support relevant experimental trials to examine the effects of education interventions (Undervisningsministeriet 2014, 5).

Norway

Number of trials

In Norway, we identified 11 RCTs initialised within the field of education (grades 1–10; pupil age 6–15). The first trial was initiated in 2009 (Buch-Iversen 2010; Elbro and Buch-Iversen 2013), and all except this first trial are still ongoing. In total, the search identified seven researcher-initiated RCTs and four policy-initiated Norwegian trials.

Funding

Most of the included studies were funded by the Research Council of Norway (5) or the Norwegian Directorate for Education and Training (4). The remaining two RCTs were funded by the University of Stavanger and University of Oslo, respectively.

Topics

The policy-initiated RCTs (4) evaluated the efficacy of governmental initiatives: for example, the increased teacher–student ratio or means taken to prevent dropout. The majority of the researcher-initiated studies tested the efficacy of different teaching methods for students with special needs or with disabilities, including students with language problems, arithmetic problems, reading difficulties and students with Down's syndrome. The only exception was an RCT that evaluated the effect of increased physical activity.

Age groups

We found interventions across grades 1–8 (pupil age 6–13), but the majority (8) of RCTs evaluated the effect of interventions given in grades 1–5 (pupil age 6–10).

Methodological preferences among policymakers

A parliamentary report from 2006 explicitly states the need for studies that test the efficacy of early interventions for vulnerable groups as well as the effect of special education (Kunnskapsdepartementet 2006). In a 2013 report to the parliament, research on how the welfare society and educational system function was defined a priority area (Kunnskapsdepartementet 2012). However, impact studies or RCTs were not explicitly mentioned. The Ministry of Education's Strategy for Educational Research 2014–2019 (Kunnskapsdepartementet 2014) and a recent report to the parliament (Kunnskapsdepartementet 2016) both explicitly highlighted impact studies. The 2016 report concluded that most of the previous evaluations of national initiatives had been

in the form of subjective evaluations and descriptions of the situation before, during and after the initiatives. The report argued that this was a limitation, but one which should be viewed in the context of (a) governmental initiatives rarely being designed in a way accustomed to impact evaluation and (b) a limited number of research groups with the competence to carry out RCTs. To overcome these problems, the report suggested that more governmental initiatives should be organised and carried out in a way that makes impact evaluation possible, meaning that the initiatives should be systematically planned and implemented, preferably assisted by researchers. The 2016 report also launched a new funding scheme for innovation in the educational sector where researchers and local school owners are encouraged to identify knowledge gaps, implement interventions and evaluate the effectiveness of the treatment (preferably by RCTs) together. The arrangement will be implemented by the Norwegian Research Council.

Sweden

Number of trials

We found 17 Swedish RCTs. The first study was initiated in 1994 (Gustafson, Samuelsson, and Jerker 2000), but the majority of studies were relatively recent. All but one trial was researcher initiated. Early RCTs were mostly small ($N < 100$), whereas all ongoing or unpublished studies have more than 100 participants.

Funding for RCTs has been provided by a diverse set of organisations, including government agencies (7), municipalities (2) and private foundations (3). Three studies have been performed without external funding, and we lack information about funding in two cases. It is noteworthy that studies have been funded by agencies such as the National Board of Health and Welfare and the Swedish Council for Working Life and Social Research, which are located organisationally under the Ministry of Social Affairs. The Ministry of Social Affairs is responsible for health care, for example, but not education. In fact, there were only two studies funded by agencies directly connected to the Ministry of Education, both of them ongoing.

Topics

Ten studies examined reading or literacy interventions aiming to improve, for example, phonological awareness, decoding, and vocabulary. Four studies evaluated mathematics interventions, and three targeted both reading and mathematics or broader skills. Nine studies examined interventions that targeted struggling or at-risk students, and eight interventions targeted more general student populations.

Age groups

We found interventions across grades 0–9 (pupil age 6–15), but the vast majority (14) was implemented in grades 0–3 (pupil age 6–9).

Methodological preferences among policymakers

The government agency mainly responsible for evaluations in grade 0–9 during the study period was the Swedish National Agency for Education. We found no discussion of the use of RCT designs on their home page. The extra search for Swedish studies revealed no

publications containing an RCT. The closest thing to a programmatic statement about RCTs was found in a large evaluation of anti-bullying methods published by the agency. There, the agency stated that although RCTs are common in evidence-based programme evaluations, it is not possible for a Swedish government agency to use this method, because it has no right to intervene in school activities and could not, for ethical reasons, require schools to use a programme with uncertain effects (Skolverket 2011, 46).

The Institute for Educational Research funded one RCT out of seven projects supported in 2016, which was the first year the institute gave out grants. In a discussion of the concept of evidence, the institute argues that it is possible to conduct RCTs within education but that other methods are also needed: '[i]t is not enough to know that something works, or rather that something has worked. We also need a deeper understanding of why it works and in what contexts' (Skolforskningsinstitutet 2017, own translation).

The recent final report of the national Swedish School Commission on how to improve learning outcomes, the quality of teaching, and promote equity in Swedish primary and secondary schools did not contain suggestions for building a stronger evidence base for Swedish education interventions through, for example, government-financed RCTs or other types of impact evaluations (Statens Offentliga Utredningar 2017).

Discussion

Our review of educational RCTs in Scandinavia shows some similar trends and tendencies across the three countries as well as some marked differences. Perhaps the most salient similarity is that few RCTs were initiated until about 2011, after which the number of trials increased in all countries, although at different rates. The historical absence of RCTs, and the subsequent break in the trend, point to some common barriers in the Scandinavian countries towards conducting RCTs. We first discuss such barriers and consider what may have provided the impetus for the change. Second, we provide some hypotheses as to why there are differences between the countries in terms of the timing and type of trials.

Common historical barriers for RCTs in Scandinavia

Research culture

Historically, educational research in all three countries has been dominated by qualitative studies. As researchers make up parts of the boards of funding agencies and serve as reviewers of proposals, a qualitatively-inclined research culture may make it difficult to obtain funding for RCTs. Educational policymakers are, furthermore, likely to be at least partly recruited among people with a background in fields related to pedagogy or educational science and to have been educated in a primarily qualitative tradition. A culture among researchers that mainly values qualitative studies may therefore extend to policymakers, who, in turn, may not push for experimental trials.

The predominance of qualitative research does not extend across all policy areas in Scandinavia. There is, for example, a much stronger tradition of using RCTs in the healthcare sector compared to the education sector in all three countries. If population size is taken into account, Denmark and Sweden are, in fact, among the countries with the most health RCTs worldwide (Torgerson and Torgerson 2008). Interestingly, as shown in the results section and in Table 2, some of the implemented and ongoing educational RCTs have been initiated or

funded by government agencies more connected to the healthcare or social sector than the educational sector (like the Swedish National Board of Health and Welfare and the Swedish Council for Working Life and Social Research, The Danish Ministry of Social Affairs and the Norwegian Institute of Public Health). We therefore believe that the dominance of qualitative research within the educational research and policy communities makes up an important explanatory factor regarding the low number of education RCTs in Scandinavia, at least historically.

The increase in RCTs during the last years may indicate a change in the culture of pedagogy or educational science. Yet several of the more recent RCTs have been initiated by researchers outside pedagogy and educational science, such as political scientists and economists. Internationally, the push for RCTs within educational research increased in the 2000s, and this has likely influenced Scandinavian researchers (Slavin 2002; Raudenbush 2005). At least in Denmark, the increasing use of RCTs was preceded by a period of an increasing number of non-RCT studies. So, in terms of research culture, RCTs may be seen as the culmination of a period of a rapidly expanding focus on impact evaluations coinciding with the introduction of exam grades in national registers. Because the countries are quite similar in this regard, we do not consider this a major explanation for the *differences* between the countries in the later years.

The recent increase in policy-initiated RCTs in Denmark and Norway does, however, point to a change in preferences in terms of RCT designs among policymakers. A contributing reason for this shift in attitudes could be the increased prevalence of assessments that facilitate international comparisons (e.g. PIRLS, PISA and TIMSS). The results from these tests have caused considerable debate in all three countries and, as the results of these comparisons in all cases were largely interpreted in negative terms, they may have provided one impetus for change. Influences from abroad may also explain some of the changes. The US, and to some extent the UK, started implementing RCTs, particularly policy-initiated RCTs, earlier than the Scandinavian countries.

Cost

RCTs are, for several reasons, often more costly than other types of experimental research designs that warrant valid causal inference, and finding the resources to run RCTs is possibly more difficult in small countries such as the Scandinavian ones. RCTs are costly because the intervention itself often involves certain costs, and, in addition, teachers and schools are often compensated financially or otherwise in RCTs, which is not a necessary expense in retrospective studies (Deding and Høg 2015). Implementation of the intervention is often closely monitored, and data collection often has to be done fully or partly in the field rather than being extracted from registers.

A substantial part of the cost of RCTs pertains to testing. There is less to gain, and it is more difficult to develop and validate tests in languages with relatively few first-language speakers. For example, the Swedish National Agency of Education had difficulty finding relevant research and expertise for the construction of test material to support the new compulsory assessment of first-grade students' reading and mathematics skills (Skolverket 2016). As mentioned in the section on school systems, all three countries now have regular national tests in place. However, these tests have sometimes been the subject of heated debates about the pros and cons of standardised testing,

which may have dissuaded researchers from using them. Another possible barrier has been that test results have either not been easily available for researchers until recently (Denmark and Norway) or that the spacing between tests has been (and is) very large (Norway and Sweden). Long spacing between tests makes test results less well suited for the short interventions often evaluated within RCTs, as post-tests either have to be made long after the intervention ended or pre-tests are not current enough.

The marked increase in educational RCTs in Denmark coincides with the implementation of national tests that: (a) are taken relatively often (voluntary tests during the autumn, and compulsory tests during the spring in many grades); (b) are available from an early age (second grade); (c) are available in several subjects including reading and mathematics; and (d) have automated elements, as the tests are taken on and are corrected by a computer. Thus, the Danish tests have important qualitative advantages and at the same time reduce the costs of performing RCTs for researchers.

Differences in the timing, prevalence, and features of RCTs between the countries

The first educational RCTs in Denmark and Sweden were conducted in the 1990s, whereas the first trial in Norway did not take place until 2009. The first Danish and Swedish studies were initiated by individual researchers or small research groups and evolved within a specific research area, dyslexia, which was apparently a very active research field and was more focussed on experimental designs at that time, compared to other education research fields. This is not surprising, as the field of special education has historically had a closer connection to both the health field and experimental psychology (Odom et al. 2005). In addition, links were also evident between the researchers involved in conducting the first studies in all three countries. Although it took place 15 years later than RCTs in Denmark and Sweden, the first RCT in Norwegian educational research also developed within the research fields of dyslexia and special education. We do not know why Norwegian researchers were not as active in this field as their Danish and Swedish colleagues.

Denmark has by far the largest number of educational RCTs compared with Norway and Sweden. It is surprising that the total number of RCTs in Sweden was only a little higher than in Norway and less than the number of trials in Denmark, considering the fact that the population of Sweden (10 million) is nearly twice as large as Denmark (5.8 million) and Norway (5.3 million), respectively. The majority of the Danish studies were initiated after 2010 with funding from the Ministry of Education. The funding for these trials was the result of a strategic push for more evidence-based knowledge and a political push for more knowledge on what works. Several trials were funded by the non-profit organisation TrygFonden. As most of the studies were conducted by researchers at TrygFonden's Centre for Child Research, TrygFonden and the Ministry of Education have played a major role in the large number of recent RCTs in educational research in Denmark.

Examining whether the trials were initiated by policymakers or researchers, we see a stark difference in the number of policy-initiated RCTs. Denmark and Norway have many policy-initiated trials within the last decade, Sweden only one. As discussed earlier, the support for RCTs among government agencies and policymakers responsible for education seems weak in Sweden compared to Denmark and Norway.

It is difficult to pin down the reasons why policymakers in Denmark and Norway have supported the use of RCTs in recent years, while there has been a lack of initiative from the

Swedish government. We mentioned earlier the debate about educational policy caused by international comparisons. Although there were some similarities in how these results were received (e.g. that the results were interpreted negatively), there were also differences. In Denmark and Norway, the results were not in line with expectations given the amount of resources used by the educational systems. This may have led to an increased focus on getting more value for money and, consequently, on getting evidence on what works. In Sweden, the political debate about causes and remedies has revolved around general features of the educational system that align well with traditional left- to right-wing political conflicts, such as segregation and the voucher system. The systemic effects of these features are difficult to study with RCTs.

Limitations

Although we performed a systematic and comprehensive search for unpublished or ongoing trials and grey literature, there may be Scandinavian educational RCTs that are not included in this paper. However, due to the nature of our search, excluded studies are likely to be ongoing rather than published studies and we therefore do not believe the addition of such studies would alter our conclusions. A more extensive qualitative approach may have contributed with a deeper understanding of the development in the three countries but was not possible due to time constraints. It is also possible that other inclusion criteria such as including trials in preschool and upper secondary education or including other research designs than RCTs could have led to other results.

Conclusion

RCTs in educational research were few in Scandinavia as a whole until about 2011, after which a marked increase in all three countries is seen, although at different rates. Up until today, the largest number of such trials has been conducted in Denmark. As the trajectories in Denmark and Norway are relatively similar, we believe that the lack of government initiatives to promote RCTs in Sweden is the most likely explanation for the differences we see across the Scandinavian countries. The large-scale funding and coordination from the government in Denmark and Norway may be crucial, which can also be gleaned from the development in considerably larger countries such as the US and the UK. The establishment of the Institute for Education Sciences in the US in 2002, and the Education Endowment Foundation in the UK in 2011, which have both been proponents of RCTs as well as provided funding for many trials, has been associated with a large increase in the number of RCTs in both countries. Government support is likely to be even more important in smaller countries such as the Scandinavian ones, where similar institutions have now been established. In light of the findings and trends discussed in this paper, we find that the use of RCT designs in Scandinavian education research is likely to continue.

Acknowledgements

The authors wish to thank those who contributed to this paper with their views on the development of Scandinavian RCTs in educational research, who helped identify relevant trials, or who provided useful comments on the manuscript. We especially appreciate the contributions from Ida

Buch-Iversen, Eva Hammar Chiriac, Carsten Elbro, Margareta Engvall, Linda Fälth, Stig Toke Gissel, Bente Eriksen Hagtvedt, Martin Hassler Hallstedt, Thomas Illum Hansen, Anders Holm, Holger Juul, Martin Karlsson, Nina Klang, Merete Konnerup, Arne Lervåg, Anita Lopez-Pedersen, Gaute Losnegard, Monica Melby-Lervåg, Frode Nyhamn, Kari-Anne Bottegaard Næss, Morten Rasmus Puck, Kari Vea Salvanes, Rune Hejlskov Schjerbeck, Jørgen Søndergård, Janne von Koss Torkildsen, Minna Törmänen, Ulrika Wolff and Björn Öckert. We are also grateful to Louise Kristiansen and Bjørn Viinholt Nielsen, who provided excellent assistance with screening and database searching.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Maiken Pontoppidan http://orcid.org/0000-0002-9491-3665
Maria Keilow http://orcid.org/0000-0003-4540-8838
Jens Dietrichson http://orcid.org/0000-0003-4178-7682
Oddny Judith Solheim http://orcid.org/0000-0002-9365-9052
Vibeke Opheim http://orcid.org/0000-0003-0060-8529
Stefan Gustafson http://orcid.org/0000-0001-9350-2955
Simon Calmar Andersen http://orcid.org/0000-0001-5274-6953

References

Andersen, S. C., L. V. Beuchert, H. S. Nielsen, and M. K. Thomsen. 2016. "The Effect of Teacher's Aides in the Classroom: Evidence from a Randomized Trial." *Unpublished manuscript*.

Andersen, S. C., M. K. Humlum, and A. B. Nandrup. 2016. "Increasing Instruction Time in School Does Increase Learning." *Proceedings of the National Academy of Sciences* 113 (27): 7481–7484. doi:10.1073/pnas.1516686113.

Andersen, S. C., and H. S. Nielsen. 2016. "Reading Intervention with a Growth Mindset Approach Improves Children's Skills." *Proceedings of the National Academy of Sciences* 113 (43): 12111–12113. doi:10.1073/pnas.1607946113.

Arnbak, E., and C. Elbro. 2000. "The Effects of Morphological Awareness Training on the Reading and Spelling Skills of Young Dyslexics." *Scandinavian Journal of Educational Research* 44 (3): 229–251. doi:10.1080/00313830050154485.

Broady, D., M. Börjesson, T. Dalberg, J. Krigh, and I. Lidegran. 2011. "Inventering Av Svensk Utbildningsvetenskaplig Forskning." *Vetenskapsrådets Rapportserie 2*. Vol. 2011. Stockholm: Vetenskapsrådet.

Buch-Iversen, I. 2010. "Betydningen av inferens for leseforståelse. Effekter av inferenstrening." PhD diss., Universitetet i Stavanger, Nasjonalt senter for leseopplæring og leseforskning.

Christensen, C. P., I. G. Andersen, P. Bingley, and C. S. Sonne-Schmidt. 2015. "Effekten Af It-Støtte På Elevers Læsefærdigheder. Et Felteksperiment I Horsens Kommune." *SFI Rapport: 14:25*. København: SFI - Det Nationale Forskningscenter for Velfærd.

Deding, M., and M. A. Høg. 2015. "Lodtrækningsforsøg Og Andre Eksperimentelle Design I Danmark." *Økonomi & Politik* 88 (3): 63–71.

Elbro, C., and E. Arnbak. 1996. "The Role of Morpheme Recognition and Morphological Awareness in Dyslexia." *Annals of Dyslexia* 46: 209–240. doi:10.1007/BF02648177.

Elbro, C., and I. Buch-Iversen. 2013. "Activation of Background Knowledge for Inference Making: Effects on Reading Comprehension." *Scientific Studies of Reading* 17 (6): 435–452. doi:10.1080/10888438.2013.774005.

Elbro, C., and L. Knudsen. 2010. "The Importance of Genre Knowledge for Text Comprehension. Insights from a Training Study." Poster presented at the Annual meeting of the Society for Text and Discourse, Chicago, USA, August 16–18.

Elbro, C., and D. K. Petersen. 2004. "Long-Term Effects of Phoneme Awareness and Letter Sound Training: An Intervention Study with Children at Risk for Dyslexia." *Journal of Educational Psychology* 96 (4): 660–670. doi:10.1037/0022-0663.96.4.660.

Elbro, C., I. Rasmussen, and B. Spelling. 1996. "Teaching Reading to Disabled Readers with Language Disorders: A Controlled Evaluation of Synthetic Speech Feedback." *Scandinavian Journal of Psychology* 37: 140–155. doi:10.1111/j.1467-9450.1996.tb00647.x.

Esping-Andersen, G., and W. Korpi. 1987. "From Poor Relief to Institutional Welfare States: The Development of Scandinavian Social Policy." In *The Scandinavian Model: Welfare States and Welfare Research*, edited by R. Erikson, E. J. Hansen, S. Ringen, and H. Uusitalo. Armonk: M. E. Sharpe.

Fälth, L., S. Gustafson, and I. Svensson. 2017. ""Phonological Awareness Training with Articulation Promotes Early Reading Development." *Education* 137 (3): 261–276.

Fälth, L., S. Gustafson, T. Tjus, M. Heimann, and I. Svensson. 2013. "Computer-Assisted Interventions Targeting Reading Skills of Children with Reading Disabilities – A Longitudinal Study." *Dyslexia* 19: 37–53. doi:10.1002/dys.1450.

Fälth, L., L. Jaensson, and K. Johansson. 2016. "Working Memory Training – A Cogmed Intervention." *International Journal of Learning, Teaching and Educational Research* 14 (2): 28–35.

Fälth, L., C. Nilvius, and E. Anvegård. 2015. "Intensive Reading with Reading Lists – An Intervention Study." *Creative Education* 6: 2403–2409. doi:10.4236/ce.2015.622246.

Gissel, S. T., and S. C. Andersen. 2017. "A Cluster-Randomized Trial Measuring the Effects of Using a Digital Learning Material that Scaffolds Decoding and Students' Reading for Meaning in Grade 2." *Unpublished manuscript*.

Gustafson, S., L. Fälth, I. Svensson, T. Tjus, and M. Heimann. 2011. "Effects of Three Interventions on the Reading Skills of Children with Reading Disabilities in Grade 2." *Journal of Learning Disabilities* 44 (2): 123–135. doi:10.1177/0022219410391187.

Gustafson, S., J. Ferreira, and J. Rönnberg. 2007. "Phonological or Orthographic Training for Children with Phonological or Orthographic Decoding Deficits." *Dyslexia* 13: 211–229. doi:10.1002/(ISSN)1099-0909.

Gustafson, S., S. Samuelsson, and R. Jerker. 2000. "Why Do Some Resist Phonological Intervention? A Swedish Longitudinal Study of Poor Readers in Grade 4." *Scandinavian Journal of Educational Research* 44: 145–162. doi:10.1080/713696666.

Have, M., J. H. Nielsen, A. K. Gejl, M. T. Ernst, K. Fredens, J. T. Støckel, N. Wedderkopp, et al. 2016. "Rationale and Design of a Randomized Controlled Trial Examining the Effect of Classroom-Based Physical Activity on Math Achievement." *BMC Public Health* 16 (1): 1–11. DOI:10.1186/s12889-016-2971-7.

Holm, A., and M. M. Jæger. 2011. "Effektmåling Af Uddannelse." *Samfundsøkonomen* 1: 18–24.

Jakobsen, M., N. Bitsch, L. B. Danielsen, and C. S. Jensen. 2016. *Elevinddragelse I Folkeskolen – Resultater Af Et Lodtrækningsforsøg I 8. Klasse*. Aarhus: Institut for Statskundskab, Aarhus Universitet.

Keilow, M., M. Friis-Hansen, S. Henze-Pedersen, and S. Ravn. 2016. "Inklusionsindsatser I Folkeskolen. Resultater Fra to Lodtrækningsforsøg." *SFI Rapport 16:16*. København: SFI - Det Nationale Forskningscenter for Velfærd.

Keilow, M., M. Friis-Hansen, R. M. Kristensen, and A. Holm. 2015. "Effekter Af Klasseledelse På Elevers Læring Og Trivsel." *SFI Rapport: 15:32*. København: SFI – Det Nationale Forskningscenter for Velfærd.

Keilow, M., A. Holm, M. Friis-Hansen, and R. M. Kristensen. 2017. "Effects of a Classroom Management Intervention on Student Concentration Performance: Results from a Randomized Controlled Trial." *Unpublished manuscript*.

Konnerup, M. 2011. ""Evidensbaseret Politik."." *Samfundsøkonomen* 1: 5–10.

Kuhnle, S., and S. Hort. 2004. *The Developmental Welfare State in Scandinavia: Lessons for the Developing World*. Geneva: UNRISD.

Kunnskapsdepartementet. 2006. … *Og Ingen Sto Igjen – Tidlig Innsats for Livslang Læring*. https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007-/id441395/sec1.

Kunnskapsdepartementet. 2012. *Lange Linjer – Kunnskap Gir Muligheter*. Kunnskapsdepartementet. https://www.regjeringen.no/no/dokumenter/meld-st-18-20122013/id716040/.

Kunnskapsdepartementet. 2014. *Kunnskapsdepartementets Strategi for Utdanningsforsking. Kvalitet Og Relevans 2014-2019*. https://www.regjeringen.no/contentassets/58266736a2f74ae9b8ffca7ad7b628d6/f-4394-n_web.pdf.

Kunnskapsdepartementet. 2016. *Lærelyst – Tidlig Innsats Og Kvalitet I Skolen*. Kunnskapsdepartementet. https://www.regjeringen.no/no/dokumenter/meld.-st.-21-20162017/id2544344/?q=lærelyst18.

Lundahl, C., and S. Tveit. 2014. "Att Legitimera Nationella Prov I Sverige Och I Norge – En Fråga Om Profession Och Tradition." *Pedagogisk Forskning I Sverige* 19 (4–5): 298–324.

Lundetræ, K., O. J. Solheim, K. Schwippert, and P. H. Uppstad. 2017. "Protocol: 'On Track', a Group-Randomized Controlled Trial of an Early Reading Intervention." *International Journal of Educational Research* 86: 87–95. doi:10.1016/j.ijer.2017.08.011.

Næss, K.-A. B., L. I. Engevik, S. Hokstad, and A. G. Mjøberg. 2017. "Vokabularstimuleringstiltak for Barn Med down Syndrom: Hva Sier Forskningen?" *Norsk Tidsskrift for Logopedi* 63 (1): 6–21.

Nemmi, F., E. Helander, O. Helenius, R. Almeida, M. Hassler, P. Räsänen, and T. Klingberg. 2016. "Behavior and Neuroimaging at Baseline Predict Individual Response to Combined Mathematical and Working Memory Training in Children." *Developmental Cognitive Neuroscience* 20: 43–51. doi:10.1016/j.dcn.2016.06.004.

Norges, F. 2013. *Midtveisrapport for Programmet UTDANNING2020. Programstyrets Egenvurdering*. Oslo: Norges forskningsråd.

Odom, S. L., E. Brantlinger, R. H. Russell Gersten, B. T. Horner, and K. R. Harris. 2005. "Research in Special Education: Scientific Methods and Evidence-Based Practices." *Exceptional Children* 71 (2): 137–148. doi:10.1177/001440290507100201.

OECD. 2016. "Education at a Glance 2016 - Indicators." http://www.oecd.org/education/skills-beyond-school/education-at-a-glance-2016-indicators.htm.

Raudenbush, S. W. 2005. "Learning from Attempts to Improve Schooling: The Contribution of Methodological Diversity." *Educational Researcher* 34 (5): 25–31. doi:10.3102/0013189X034005025.

Resaland, G. K., E. Aadland, V. F. Moe, K. N. Aadland, T. Skrede, M. Stavnsbo, L. Suominen, et al. 2016. "Effects of Physical Activity on Schoolchildren's Academic Performance: The Active Smarter Kids (ASK) Cluster-Randomized Controlled Trial". *Preventive Medicine* 91: 322–328. doi:10.1016/j.ypmed.2016.09.005.

Resaland, G. K., V. F. Moe, E. Aadland, J. Steene-Johannessen, J. R. Øyvind Glosvik, O. M. Andersen, H. A. M. Kvalheim, and S. A. Anderssen. 2015. "Active Smarter Kids (ASK): Rationale and Design of a Cluster-Randomized Controlled Trial Investigating the Effects of Daily Physical Activity on Children's Academic Performance and Risk Factors for Non-Communicable Diseases." *BMC Public Health* 15 (1): 1–10. doi:10.1186/s12889-015-2049-y.

Skolforskningsinstitutet. 2017. "Evidens Betyder Stöd Eller Belägg – Inte Bevis." Skolforskningsinstitutet. Accessed September 3, 2017. http://www.skolfi.se/forskningssammanstallningar/nyckelbegrepp/evidens-betyder-stod-eller-belagg-inte-bevis/.

Skolverket. 2011. "Utvärdering Av Metoder Mot Mobbning." *Rapport 353*. Stockholm: Skolverket.

Skolverket. 2016. "Redovisning Av Uppdrag Om Kunskapskrav I Läsförståelse Och Obligatoriska Bedömningsstöd I Årskurs 1." In *Redovisning Av Regeringsuppdrag, Dnr U2015/03529/S*. Stockholm: Skolverket.

Skolverket. 2017a. "Nationella Prov I Skolan." Skolverket. Accessed September 3, 2017. https://www.skolverket.se/bedomning/nationella-prov/alla-nationella-prov-i-skolan

Skolverket. 2017b. "Provresultat I Grundskolan." Skolverket. Accessed September 3. https://www.skolverket.se/statistik-och-utvardering/provresultat/provresultat-i-grundskolan-1.218122.

Skolverket. 2017c. "Statistik Om Förskoleklass." Skolverket. Accessed September 3. https://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/forskoleklass.

Skolverket. 2018. "Nationella Prov". Skolverket. Accessed March 2, 2018. https://www.skolverket.se/bedomning/nationella-prov

Slavin, R. E. 2002. "Evidence-Based Education Policies: Transforming Educational Practice and Research." *Educational Researcher* 31 (7): 15–21. doi:10.3102/0013189X031007015.

Solheim, O. J., J. C. Frijters, K. Lundetræ, and P. H. Uppstad, 2018. "Effectiveness of an Early Reading Intervention in a Semi-Transparent Orthography: A Group Randomised Controlled Trial." *Unpublished manuscript*.

Solheim, O. J., M. Rege, and M. Erin. 2017. "Study Protocol: 'Two Teachers' a Randomized Controlled Trial Investigating Individual and Complementary Effects of Teacher-Student Ratio in Literacy Instruction and Professional Development for Teachers." *International Journal of Educational Research* 86: 122–130. doi:10.1016/j.ijer.2017.09.002.

Statens Offentliga Utredningar, 2013. *Friskolorna I Samhället*. Vol. 56. SOU. Stockholm: Utbildningsdepartmentet

Statens Offentliga Utredningar, 2017. *Samling För Skolan. Nationell Strategi För Kunskap Och Likvärdighet*. Vol. 35. SOU. Stockholm: Utbildningsdepartementet

Sterner, G., U. Wolff, and O. Helenius. 2015. *Tal, Resonemang Och Representationer - En Interventionsstudie I Matematik I Förskoleklass*. Licentiatuppsats: University of Gothenburg.

Sundell, K., and E. Stensson. 2010. *Effektutvärderingar I Doktorsavhandlingar*. Stockholm: Socialstyrelsen. http://www.socialstyrelsen.se/publikationer2010/2010-3-24.

Tarp, J., S. L. Domazet, C. H. Karsten Froberg, H. L. B. Andersen, and A. Bugge. 2016. "Effectiveness of A School-Based Physical Activity Intervention on Cognitive Performance in Danish Adolescents: LCoMotion—Learning, Cognition and Motion – A Cluster Randomized Controlled Trial." *PLoS ONE* 11 (6): 1–19. doi:10.1371/journal.pone.0158087.

Torgerson, D. J., and C. J. Torgerson. 2008. *Designing Randomised Trials in Health, Education and the Social Sciences. An Introduction*. Basingstoke: Palgrave Macmillan.

Törmänen, M., and M. Takala. 2009. "Auditory Processing in Developmental Dyslexia: An Exploratory Study of an Auditory and Visual Matching Training Program with Swedish Children with Developmental Dyslexia." *Scandinavian Journal of Psychology* 50 (3): 277–285. doi:10.1111/j.1467-9450.2009.00708.x.

Undervisningsministeriet. 2014. *Strategi for Forsøg Og Udvikling I Undervisningsministeriet*. København: Uddannelsesministeriet.

Undervisningsministeriet. 2017a. "Om De Nationale Test". Undervisningsministeriet. Accessed March 2, 2018. https://www.uvm.dk/folkeskolen/elevplaner-nationale-test-og-trivselsmaaling/nationale-test/om-de-nationale-test

Undervisningsministeriet. 2017b. *Vejledning Om De Nationale Test – Til Kommuner*. København: Uddannelsesministeriet.

UNESCO Institute of Statistics. 2012. *International Standard Classification of Education. ISCED 2011. Montreal*. Canada: UNESCO Institute for Statistics.

Utdanningsdirektoratet. 2017. "Kva Er Nasjonale Prøver?" Utdanningsdirektoratet. Accessed September 3, 2017. https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/om-nasjonale-prover/.

Viby Mogensen, G. 2010. *Det Danske Velfærdssamfunds Historie: Tiden Efter 1970. Bind 1*. København: Gyldendal.

Wolff, U. 2011. "Effects of a Randomised Reading Intervention Study: An Application of Structural Equation Modelling." *Dyslexia* 17: 295–311. doi:10.1002/dys.v17.4.

Wolff, U. 2016. "Effects of a Randomized Reading Intervention Study Aimed at 9-Year-Olds: A 5-Year Follow-Up." *Dyslexia* 22: 85–100. doi:10.1002/dys.v22.2.