

En multisensorisk läsinlärningsmetod för elever med risk för lässvårigheter

Resultat från ett randomiserat
fältexperiment

Martin Bøg

Jens Dietrichson

Anna Aldenius Isaksson

En multisensorisk läsinlärningsmetod för elever med risk för lässvårigheter^a

Resultat från ett randomiserat fältexperiment

av

Martin Bøg^b, Jens Dietrichson^c och Anna Aldenius Isaksson^d

2019-04-15

Sammanfattning

Trots att läsning är en grundläggande färdighet lämnar många elever grundskolan med stora brister i läsförståelse. Vi undersöker effekterna av en läsinlärningsmetod som riktar sig till de elever i förskoleklass och årskurs ett som har störst risk att hamna i lässvårigheter. Metoden bygger på multisensoriska inlärningsmetoder, fokuserar på fonologisk medvetenhet och kopplingen mellan ljud och bokstäver och genomförs i små grupper med en lärare och en eller två elever. Vi använder ett randomiserat fältexperiment med 161 elever på 12 svenska skolor för att analysera effekterna av metoden och finner stora positiva effekter på våra två primära utfallsmått: ett standardiserat test av avkodningsförmåga och ett standardiserat test av bokstavskännet. Vi finner också positiva effekter på ett test av fonologisk medvetenhet och ett mått på självtillit, samt små och statistiskt insignifikanta effekter på mått på läsglädje och motivation. Metoden står sig väl jämfört med liknande metoder i termer av kostnadseffektivitet.

^a Rapporten är en populärvetenskaplig sammanfattning av Bøg m.fl. (2019), IFAU Working Paper 2019:7. Vi vill tacka medverkande skolor, Åsa Malmberg och Christina Rick utan vilka studien inte hade kunnat genomföras, samt Trine Filges, Caroline Hall, Mette Kjærgaard Thomsen, Björn Öckert och seminariedeltagare vid TrykFondens Børneforskningscenter, IFAU och CEN, vars kommentarer förbättrade rapporten. IFAU:s forskningsbidrag 198/2016 möjliggjorde studien. Intresseförklaring: Isaksson har utvecklat den metod som studeras och tillhandahåller den kommersiellt till skolor. Analysen genomfördes av Bøg och Dietrichson.

^b Bøg: Lundbeck A/S, Ottiliavej 9, DK 2500, Valby, Denmark; martin.bog@gmail.com

^c Dietrichson. Kontaktperson för rapporten. VIVE–The Danish Center for Social Science Research, Copenhagen, Denmark. jsd@vive.dk

^d Isaksson, S:t Jörgens plats 8, 252 20 Helsingborg, Sweden. anna@lararkonsult.se

Innehållsförteckning

1	Inledning.....	3
1.1	Läsklar.....	6
2	Metoder.....	8
2.1	Utfallsmått.....	8
2.2	Deltagare.....	10
2.3	Kontrollgruppens undervisning.....	15
2.4	Implementering.....	15
2.5	Analysplan.....	16
3	Resultat.....	19
3.1	Kostnader och kostnadseffektivitet.....	21
3.2	Resultat vid uppföljningsmätningen.....	22
4	Diskussion.....	23
4.1	Perspektiv på effekter och kostnadseffektivitet.....	23
4.2	Studiedesign och effekter.....	25
4.3	Begränsningar.....	26
5	Slutsatser.....	28
6	Källor.....	29

1 Inledning

Att kunna läsa väl är en grundläggande färdighet och läsförmåga är starkt korrelerad med mått på framgång på arbetsmarknaden (t.ex. OECD, 2016a). I genomsnitt bedöms cirka 20 procent av 15-åringarna i OECD-länderna ha betydande brister när det gäller läsförståelse och alla länder har en stor grupp elever som läser på en nivå som innebär att de har svårt att tillgodogöra sig information från mer komplexa texter (OECD, 2016b). Andelen med bristande läsförståelse är vidare ungefär lika stor bland vuxna (OECD, 2016a). Att hitta kostnadseffektiva metoder för att hjälpa elever som riskerar att hamna i lässvårigheter bör därför vara en central uppgift för forskare och beslutsfattare i skolan.

Den tidigare litteraturen ger skäl till att vara optimistisk om att det går att förbättra denna situation. Forskningsöversikter som jämför effekterna av riktade (eller selekterade/indikerade) insatser för riskgrupper tyder på att vissa typer av insatser har betydande kortsiktiga effekter (t.ex. Slavin m.fl., 2011; Dietrichson m.fl., 2017). I jämförelser mellan undervisningsmetoder är de genomsnittliga effekterna vanligen störst för ”tutoring”, det vill säga intensivundervisning i små grupper med vuxna tutors (Slavin m.fl., 2011; Dietrichson m.fl., 2017). När det gäller innehållet i insatser är kopplingen mellan ljud och bokstäver och fonologisk medvetenhet viktiga byggstenar för barns tidiga läsinläring (t.ex. Bus & van IJzendoorn, 1999; Snow m.fl., 1998; Ehri m.fl., 2001; Snowling & Hulme, 2011; Statens beredning för medicinsk utvärdering, 2014), särskilt för barn som riskerar att hamna i lässvårigheter (t.ex. Lundberg & Høien, 1996; Machin m.fl., 2018). Fonologisk medvetenhet innebär att barn kan urskilja fonemen, de ljudenheter som är de minsta byggstenarna i språket, och kombinera dem till större enheter, såsom stavelser.

Viktig kunskap saknas dock. Effekterna varierar stort mellan olika insatser, även inom kategorier av insatser som använder liknande metoder och innehåll (t.ex. Dietrichson m.fl., 2017). Det finns alltså mycket kvar att lära när det gäller hur effektiva insatser ska utformas. Dessutom har få studier utvärderat de långsiktiga effekterna av riktade interventioner och undersökt kostnadseffektiviteten. Eftersom de kortsiktiga effekterna av riktade insatser är ungefär lika stora genom grundskolan är den optimala tidpunkten för insatser också en öppen fråga (se t.ex. Hill m.fl., 2008; Dietrichson m.fl., 2017, 2019).

Vi använder ett randomiserat fältexperiment där 161 elever i 12 svenska skolor deltog för att undersöka de kortsiktiga effekterna av en läsinlärningsmetod – *Läsklar* – utvecklad av den tredje författaren till denna rapport. Metoden, som riktar sig till de elever i förskoleklass och årskurs 1 med störst risk att hamna i lässvårigheter, kombinerar tre huvudkomponenter: (i) multisensoriska metoder

som syftar till att aktivera många sinnen och därigenom förbättra minneskapaciteten; (ii) ett fokus på fonologisk medvetenhet och kopplingen mellan bokstäver och ljud; och (iii) intensivundervisning en-till-en eller en-till-två av lärare (d.v.s. tutoring).

Eleverna tränar med metoden redan i förskoleklass och årskurs 1, vilket kan ha minst tre viktiga fördelar. För det första så är elever som inte kan läsa efter årskurs 1 starkt överrepresenterade bland elever med lässvårigheter senare i grundskolan (t.ex. Francis m.fl., 1996; Denton m.fl., 2004). Tidiga insatser kan förhindra att elever behöver uppleva långa perioder av skolmisslyckande.

För det andra finns det studier som tyder på att barn börjar se kopplingen mellan prestationer och den egna förmågan när de är ungefär 5–7 år gamla, det vill säga ungefär när de börjar grundskolan (t.ex. Nicholls, 1978; Butler, 1999; Muenks & Miele, 2017). Vidare befinner sig barn före starten av grundskolan sällan i situationer där deras prestationer systematiskt jämförs med deras kamrater (Poskiparta m.fl., 2003). Länsatser i förskoleklass och årskurs 1 kan därför lyfta eleverna innan de börjar se sig själva som svaga läsare.

För det tredje så kan senare insatser behöva vara mer omfattande och dyrare om de ska kunna avhjälpa både lässvårigheter och motivationsproblem (Cook m.fl., 2014; Vaughn m.fl., 2015). Om utvecklingen av senare färdigheter bygger på tidigare färdigheter kan framgångsrika tidiga insatser vidare vara viktiga för resultatet av senare insatser (t.ex. Stanovich, 1986, Cunha & Heckman, 2007). Tidiga insatser kan därför vara mer kostnadseffektiva.

Vi finner positiva och statistiskt signifikanta effekter på våra två primära utfallsmått, ett standardiserat test av avkodning och ett standardiserat test av bokstavskänedom. Vi finner också positiva och signifikanta effekter på test av fonologisk medvetenhet och ett mått på elevens tro på sin egen förmåga. Effekterna är små och inte statistiskt signifikant skilda från noll på mått på läsglädje och motivation, men nästan alla elever är starkt motiverade och tycker att det ska bli kul att lära sig att läsa vid samtliga mätningar. Dessa resultat stöder tanken att vi genomförde insatsen innan eleverna tappat motivationen för läsning.

De statistiskt signifikanta effekterna är stora. De är exempelvis två till fyra gånger större än de 0,25 standardavvikelser som anses vara betydande effekter av What Works Clearinghouse (2014). Effekterna är också stora jämfört med de genomsnittliga effekterna av andra insatser, inklusive liknande tutoring-insatser. Metoden är dessutom kostnadseffektiv i jämförelse med liknande insatser. Vi diskuterar dessa jämförelser mer ingående i den avslutande diskussionen.

Vår studie bidrar till litteraturen om riktade länsatser för elever med risk att hamna i lässvårigheter på följande sätt: *Läsklar* adderar den multisensoriska

metoden till de mer välstuderade insatskomponenterna tutoring, fonologisk medvetenhet och bokstav-ljudkopplingen. Neurovetenskapliga studier tyder på att multisensorisk stimulans aktiverar fler inlärningsmekanismer i hjärnan än att bara ett sinne stimuleras (Shams & Seitz 2008; Shams m.fl., 2011) och barn är bättre på kategorisering och numeriska matchningsuppgifter efter att ha fått multisensorisk stimulans (Jordan & Baker, 2011; Broadbent m.fl., 2018). Multisensoriska metoder har länge ansetts vara god praxis för att lära elever med dyslexi att läsa (Snowling & Hulme, 2011), men systematiska studier som jämför multisensoriska metoder med andra metoder för elever med risk att hamna i lässvårigheter är fortfarande relativt få (Broadbent m.fl., 2018; Snowling & Hulme, 2011).¹

Vi genomförde insatsen på ett sätt som liknar hur skolor normalt skulle genomföra metoden. Vi övervakade inte hur skolorna genomförde insatsen och stöttade inte skolorna mer än vad som skulle ha varit fallet utanför studien. Vidare exkluderade vi inte elever med de allra största problemen, utan försökte aktivt inkludera dem, oavsett orsaken till deras problem. Detta är också i linje med hur skolor brukar arbeta, medan forskningsstudier ofta exkluderar elever som till exempel håller på att lära sig språket i landet, har neuropsykiatriska diagnoser och andra funktionsvariationer, eller har de allra lägsta resultaten på tester av läsförmåga (t.ex. Amendum m.fl., 2011; Burns m.fl., 2003; Lovett m.fl., 2017; Vadasy m.fl., 2006).

Vi undersöker också kostnadseffektiviteten av insatsen och testar effekterna på elevernas självrapporterade attityder till läsning, vilket det finns få andra exempel på i tidigare litteratur.² Tutoring där lärare genomför insatsen är i allmänhet en resursintensiv insats, men vi visar att det är möjligt att väsentligt förbättra elevernas resultat med relativt små medel och med hjälp av en kort men intensiv insats.

Slutligen, i en litteratur dominerad av studier från engelsktalande länder, i synnerhet USA, tillhandahåller vi resultat från en annan kontext. Det finns ingen garanti för att insatser som fungerar bra i ett sammanhang också fungerar bra i andra. Vi känner bara till tre tidigare studier från de nordiska länderna om effekterna av läsinsatser som riktar sig specifikt till elever i förskoleklass och årskurs 1 med risk att hamna i lässvårigheter. Vår egen pilotstudie av *Läsklar* använde en kvasi-experimentell design för att undersöka effekterna i tre svenska skolor. För en jämförbar riskgrupp hittade vi liknande korttidseffekter på

¹ Se Sadoski & Willson (2006) och Ehri et al. (2007) för exempel på framgångsrika läsinterventioner som använder multisensoriska metoder.

² För undantag när det gäller kostnadseffektivitet, se Hollands m.fl. (2013, 2016) och Jacob m.fl. (2016). Morgan m.fl. (2008) och Fives m.fl. (2013) använder mått på attityder, där elever i årskurs 1 får svara själva.

avkodning och mindre effekter på bokstavskännetid (Bøg m.fl., 2018). Elbro och Petersen (2004) finner positiva effekter sju år efter en insats som syftade till att förbättra den fonologiska medvetenheten hos danska förskoleklass elever med dyslektiska föräldrar. Rogde m.fl. (2016) hittar positiva kortsiktiga effekter av en insats där barn med annat modersmål i norska förskoleklasser fick en träna språkkunskaper i små grupper.³

1.1 Läsklar

Läsklar har tre huvudsyften. Det första är att hjälpa eleverna förstå att ord dels anger ett innehåll och dels har en form som består av fonem. Det andra är att hjälpa eleverna förstå hur bokstäver och ljud hänger ihop. Det tredje är att visa hur man avkodar enkla ord, det vill säga att förstå hur man gör när man läser ihop bokstäver till ord. I ”the simple view of reading” är effektiv avkodning en nödvändig komponent för läsförståelse (t.ex. Hoover & Tunmer, 2018).

Eleverna tränade tillsammans med en lärare, antingen individuellt eller i grupper om två elever (16 elever tränade i grupper om två). Intentionen var att eleverna skulle träna 3–4 gånger i veckan under 8–10 veckor, i totalt 30–35 sessioner där varje session varar i ungefär 10–15 minuter. I genomsnitt fick eleverna i interventionsgruppen 32,6 sessioner, medianen var 32 och minimum och maximum var 21 och 50.

Lärarna fick utbildning i metoden genom en videokurs, som består av korta filmer där den tredje författaren visar hur sessionerna går till. Den tredje författaren besökte också varje skola en eller två gånger för att visa hur man genomför metoden och svara på frågor.

Metoden innehöll tre typer av sessioner eller steg. I steg 1 används 25 figurer gjorda i cernitlera med begynnelsebokstäver som motsvarar 25 av de svenska språkljuden, en låda med 29 fack – ett fack för varje bokstav – samt ett inplastat ark som visar handalfabetet. Till första sessionen har läraren valt ut tre figurer och övningen går ut på att placera figuren i sitt fack eller sitt ”hus”. För att figuren ska få komma in i sitt hus ska eleven lista ut en kod, det vill säga lista ut hur det första bokstavsljudet i figurens namn låter. Till exempel ”ll” som i lejon. När eleven listat ut att lejon börjar på l så ska eleven säga ”ll” och samtidigt göra ett ”l” med handalfabetet. Sen kan figuren flytta in i sitt hus. När alla tre figurerna har fått flytta in i sitt hus, så repeterar man en gång och så är sessionen slut.

³ Fälth m.fl. (2017) utvärderar en helklassinsats som fokuserade på träning av fonologisk medvetenhet i fyra svenska förskoleklasser. Insatsen riktade sig inte specifikt mot riskgrupper, men studien finner positiva effekter också för denna grupp. Wolff (2011, 2016) undersöker en tutoring-insats med fonologisk träning för elever med lässvårigheter i tredje klass. Insatsen resulterade i signifikant förbättrade läsförmågor på flera områden direkt efter interventionen och på avkodningsförmåga fem år efter interventionen.

Denna del tar ungefär 10–15 minuter. Vid nästa session får eleven först flytta in med de tidigare avklarade figurerna och sedan öka på med tre nya figurer.

När eleverna kan cirka 15 begynnelsebokstäver, vilket brukar ta ungefär en månad, är det dags att börja med steg 2. I steg 2 används bilder som representerar enkla ord och är klistrade på ett vitt papper. Steg 2 innebär att man ska lyssna ut alla ljud i orden som är på tre av bilderna i kopieringsunderlaget. Bilderna klistras på baksidan av pappret. Läraren agerar sekreterare till eleverna när de ljudar fram vilka ljudordet innehåller, varefter läraren skriver dessa på framsidan. När detta är klart sätts ordkorten in i en pärm där framsidan blir det skrivna ordet och baksidan blir en bild av ordet. Det gör att eleverna kan öva själva, till exempel hemma, eftersom de kan kontrollera om de har läst rätt genom att vända på sidan. Pärmen fylls successivt på tills den innehåller ungefär 25 ord. Om det behövs mer övning, binds de gamla orden ihop till en bok och sedan påbörjas 25 nya ord.

Den tredje typen av session syftar till att bygga hastighet och flyt i avkodningen genom att öva med en kortlek där korten visar korta stavelser (elever som inte lärt sig avkoda genomför därför inte den här övningen). Övningen går ut på att avkoda så många stavelser som möjligt under en minut. Läraren registrerar resultaten i ett diagram och sedan repeterar man med samma stavelser två gånger till i samma session. Detta ger i regel en tydlig förbättring mellan gångerna som eleverna kan följa i diagrammet. Att göra förbättringarna synliga för eleverna motiverar dem ofta till att öva mera.

Idén bakom *Läsklar* är att fokusera på fonologisk medvetenhet och kopplingen mellan ljud och bokstäver och därigenom hjälpa eleverna förstå att ljudströmmen som produceras när man talar kan delas in i ett antal bokstavsljud och att man kan skilja dessa åt. De små undervisningsgrupperna tillåter individualiserat stöd och anpassad hastighet, samt ger möjlighet till tät återkoppling från läraren. Tanken med den multisensoriska träningen är att involvera så många sinnen som möjligt och därmed aktivera hjärnans inlärningsmekanismer och stärka minneskapaciteten. När eleverna hjälper figurerna att flytta in i sina rum får minnet en extra hävstång genom att det kinestetiska, taktila, visuella, och auditiva sinnet aktiveras. Det skapas också en liten berättelse kring figurerna som flyttar in i huset, vilket kan göra att det episodiska minnet aktiveras (minnet för självbiografiska händelser, se t.ex. Tulving, 2002). Genom att använda handalfabetet aktiveras också flera sinnen, samtidigt som eleverna får med sig detta redskap som minnehjälp när de går över till den ordinarie textläsningen. Att jobba med figurerna är också ett sätt att stärka elevernas motivation. Tanken är att det är lättare och roligare för eleverna att lista ut var figurerna ska bo och hjälpa dem in i sina nya hem, än att enbart fundera på hur ord låter i början.

2 Metoder

I det här avsnittet beskrivs utfallsmåtten, deltagarna, hur vi delade in eleverna i interventions- och kontrollgrupper, kontrollgruppens undervisning, implementeringen och analysplanen. Vår engelska rapport (Bøg m.fl., 2019) innehåller fler detaljer om metoderna.

2.1 Utfallsmått

Våra primära utfallsmått, ett standardiserat test av avkodningsförmåga och ett standardiserat test av hur många bokstäver eleverna kan, ligger nära syftet med *Läsklar*. Som ett sekundärt utfallsmått använder vi ett test av fonologisk medvetenhet, utvecklat av den tredje författaren. För att undersöka om programmet påverkar attityder till läsning, använder vi tre mått relaterade till självtillit, läsglädje och motivation.

Vi genomförde en testomgång innan eleverna fördelades mellan interventions- och kontrollgruppen (*för-mätning*) och igen kort efter att interventionsgruppen hade fått träna med metoden (*efter-mätning*) och sedan en tredje gång efter att kontrollgruppen också tränat med metoden (*uppföljnings-mätning*). Skolor som genomförde metoden i förskoleklass kunde börja både under höst- och vårtermin och vi genomförde för-mätningar mellan slutet av augusti och början av mars. Alla skolor som genomförde metoden i första klass började tidigt i årskurs 1 och vi genomförde för-mätningar från slutet av augusti till början av oktober. Det genomsnittliga antalet dagar mellan för- och efter-mätningarna är 103 och det genomsnittliga antalet dagar mellan efter- och uppföljningsmätningarna är 102.

Nästan alla elever testades av samma testledare (en tidigare lärare), som inte visste om vilka elever som hade tränat med *Läsklar*. I en testomgång på en skola använde vi en annan testledare, som heller inte visste vilka elever som tränat med metoden. Testledarna genomförde mätningarna individuellt med eleverna, vilket tog ungefär 10–15 min. Den totala poängen för varje utfallsmått, tillfälle och elev plus anonymiserade elevidentifikationsnummer överfördes till en Excel-fil och skickades till forskarna.

2.1.1 Avkodning

LäSt är ett testbatteri som används för identifiering och diagnos av grundläggande läsfärdigheter (Elwér m.fl., 2016). Materialet innehåller delprov av avkodningsförmåga, stavning och läsförståelse. Vi använder delprovet av avkodningsförmåga som har två delar (Avkodning ord, del A och B) där eleverna läser korta

enskilda ord (riktiga ord, inte pseudo-ord) med ökande svårigheter i 45 sekunder i varje del, och poäng ges för varje korrekt ord (maxpoäng är 200).

Test-re-test-tillförlitligheten hos både del A och B i avkodningstestet är 0,93 och korrelationerna mellan de två delarna och två andra vanliga svenska test (*Ordkedjor* och *H4*) ligger mellan 0,59 och 0,88 (Elwér m.fl., 2016). Testkonstruktionen liknar det i amerikanska studier ofta använda *Word Reading Efficiency*, som ingår som ett delprov i *Test of Word Reading Efficiency* (Torgesen m.fl., 1999), där eleverna också ska försöka att läsa isolerade ord med ökande svårighet under 45 sekunder.

2.1.2 Bokstavskänedom

Vi använder ett diagnostiskt prov från testmaterialet *LäsEttan* som testar hur eleverna klarar att snabbt benämna bokstäverna i alfabetet. Delprovet tar ungefär 1–2 min att genomföra. Maxpoäng är 27, en för varje bokstav som ingår i testet (Johansson, 2009). Information om tillförlitlighet och korrelationer med andra test saknas i testmaterialet. Testet liknar t.ex. delprovet *Letter Naming Fluency* som ingår i *Dynamic Indicators of Basic Literacy Skills* (se t.ex. Good et al., 2002), ett vanligt testbatteri i amerikanska studier.

2.1.3 Fonologisk medvetenhet

Vi använder ett kort test av fonologisk medvetenhet som undersöker om eleverna känner till det första ljudet i 10 ord representerade av en bild på ett objekt eller ett djur. Eleven ska först identifiera det avbildade djuret eller objektet och sedan säga det ljud ordet börjar med. Varje korrekt svar är värt en poäng och maxpoängen är därför 10. Det här testet har utvecklats av den tredje författaren för denna studie med det primära syftet att urskilja de elever som behöver extra hjälp. Testet har en låg maxpoäng och eftersom den ordinarie undervisningen i svenska skolor ofta fokuserar på fonologisk medvetenhet kan det, särskilt i årskurs 1, ha betydande takeffekter.⁴ Av dessa skäl är detta test ett sekundärt utfall.

⁴ Takeffekter uppstår när elever når maxpoängen på ett test, vilket gör att inga förbättringar upptäcks även om den underliggande förmågan som testet mäter har förbättrats ytterligare. När elever förbättrar sin förmåga men utan att nå minimumpoängen på ett test kallas det golveffekter, vilket igen gör att förbättringen inte upptäcks.

2.1.4 Attityder till läsning

Vi använder tre enkla frågor som ställs till de deltagande eleverna för att mäta effekter på attityder till läsning relaterade till självtillit, läsglädje och motivation:

- Hur lätt eller svårt är det att lära sig läsa, tror du?
- Hur roligt tycker du att ska bli att lära sig läsa?
- Hur gärna vill du lära dig läsa?

För varje fråga läste testledaren frågan och eleverna angav sina svar på en visuell analog skala (VAS), där de två ändpunkterna visades med hjälp av en ledsen och en glad gubbe. Testledaren förklarade att den ledsna gubben till exempel betydde att det inte blir roligt att lära sig läsa och den glada gubben att det ska bli mycket roligt. Elevernas svar omvandlas sedan till en skala med lägsta poäng 0 och högsta 10.

Vi känner inte till några svenska validerade mått på attityder till läsning där elever i förskoleklass och årskurs 1 svarar själva, men liknande tester har tidigare använts i exempelvis Tideman m.fl. (2011). Vi ville använda frågor där eleverna svarar själva eftersom lärarna vet vilka elever som tränat med metoden och lärarnas bedömning av elevernas attityder kan därmed vara svåra att separera från lärarnas uppfattning om metoden.

2.2 Deltagare

2.2.1 Skolor och lärare

Vi rekryterade 12 skolor från 7 kommuner i Skåne under hösten 2016 och våren 2017. De sex första skolorna började genomföra *Läsklar* i förskoleklass i slutet av 2016 eller under våren 2017. De sex skolor som rekryterades sist började genomförandet i förskoleklass eller årskurs 1 under hösten 2017. Fyra skolor genomförde en andra omgång av metoden för en ny årskull elever. Totalt genomfördes därmed 16 omgångar. De sista eleverna avslutade sin träning med metoden i början av juni 2018.

Tabell 1 jämför de deltagande skolorna med det nationella genomsnittet av ett antal socioekonomiska variabler, prov- och betygsresultat, samt visar beskrivande statistik om de deltagande lärarnas egenskaper. De deltagande skolorna har en lägre andel föräldrar med högre utbildning, högre andel elever med utländsk bakgrund, en lägre andel elever som har godkänt i alla ämnen i årskurs 9 och ett lägre medelvärde på de nationella testen i svenska och matematik i årskurs 6. Det finns emellertid variation mellan skolorna, vilket framgår av det relativt stora intervallet mellan minimumvärdet i kolumn (3) och maximumvärdet i kolumn (4).

Totalt deltog 23 lärare, specialpedagoger och speciallärare i studien som tutors. De var i regel erfarna lärare (de har i genomsnitt arbetat i 23 år som lärare), men bara ungefär en tredjedel hade tidigare arbetat med liknande metoder (inklusive tre som arbetade med *Läsklar* i vår pilotstudie).

2.2.2 Elever

Totalt deltog 161 elever i antingen interventions- eller kontrollgruppen (se avsnitt 2.2.3 för en beskrivning av hur vi fördelade elever mellan grupperna), 130 i förskoleklass och 31 i årskurs 1. Vi analyserar dem tillsammans i analysen.

Tabell 1 Beskrivande statistik för deltagande skolor och lärare

Variabler	(1) Nationellt genomsnitt	(2) Urvals- genomsnitt	(3) Min	(4) Max	N
Skolor					
Föräldrar med högre utbildning	57%	40%	29%	60%	12
Utländsk bakgrund	24%	52%	11%	91%	12
Godkänt i alla ämnen i årskurs 9	74%	56%	33%	73%	6
Nationellt prov i matematik i årskurs 6	12,1	10,4	6,4	13,0	11
Nationellt prov i svenska i årskurs 6	13,0	12,3	8,5	15,1	10
Lärare					
Lärarerfarenhet (år)	-	23	8	42	21
Tidigare erfarenhet med liknande metoder	-	33%	0	1	21
Speciallärare/Specialpedagog	-	45%	0	1	21

Not: Källa för skolvariablerna: Skolverket (2018a). Statistiken gäller för skolåret 2016/2017 och för F-6 delen av varje skola, när denna information var tillgänglig/relevant. Utländsk bakgrund definieras som att antingen vara född utanför Sverige eller ha två föräldrar som är födda utanför Sverige. Notera att för variablerna godkänt i alla ämnen i årskurs 9, nationellt prov i matematik i årskurs 6 och nationellt prov i svenska i årskurs 6 så saknar vi information för några skolor, t.ex. för att inte alla skolor har årskurser efter årskurs 6, eller har för få elever som gör ett visst prov. Källa för lärarvariablerna: intervjuer och enkät till medverkande lärare. Vi saknar information från två lärare. N betecknar antalet skolor eller lärare.

I analysen fokuserar vi på gruppen elever med kompletta resultat på både för- och efter-mätningen. Få elever lämnade studien: tre elever flyttade till andra skolor efter randomiseringen (en från interventionsgruppen och två från kontrollgruppen), en interventionsgruppselev hade inte fullständiga för-mättningsresultat och en kontrollgruppselev var frånvarande vid efter-mätningen (men närvarande vid uppföljningsmätningen). Sexton ytterligare elever, jämnt fördelade mellan grupperna, lämnade studien mellan efter- och uppföljningsmätningen.

Ungefär 45 procent är flickor, och 20 procent är elever i årskurs 1. Deltagande elever var de som skolorna ansåg hade störst risk att hamna i lässvårigheter. För de flesta elever specificerade inte skolorna varför de hade en förhöjd risk, men vårt urval inkluderar nyanlända elever, elever med beteendeproblem (t.ex. ADHD eller koncentrationssvårigheter), språkstörningar, hörselproblem, kognitiva funktionsnedsättningar, eller misstänkt dyslexi. Totalt 24 procent av urvalet hade någon form av specificerad orsak till varför de hade en förhöjd risk för lässvårigheter.

2.2.3 Fördelning av elever till interventions- och kontrollgrupp

Vår studie är ett randomiserat fältexperiment, där vi fördelade eleverna till interventions- och kontrollgruppen slumpmässigt. Vi valde en väntelista-design, där kontrollgruppen får insatsen efter interventionsgruppen, av främst två skäl. För det första, alla elever fick chansen att träna med *Läsklar*. Detta gjorde det sannolikt också enklare att rekrytera skolor och gjorde dem mer villiga att acceptera randomiseringen. För det andra kan en kontrollgrupp (eller deras föräldrar och lärare), som vet att den kommer att få en insats vid ett senare tillfälle, vara mindre benägen att söka extra hjälp någon annanstans jämfört med en grupp som vet att den inte får någonting. Därmed minskar risken för kontrollgruppen också påverkas av insatsen samtidigt med interventionsgruppen.

Vi använde en fördelningsprocedur med tre steg där eleverna fördelades separat inom varje deltagande skola. Vi tillämpade proceduren sekventiellt, det vill säga så snart en skola sagt ja till att delta, så genomförde vi följande steg:

Steg 1: Lärare och specialpedagoger i varje skola valde ut den grupp elever som de trodde var mest i behov av extra hjälp, oavsett orsaken till deras svårigheter. Denna grupp testades sedan med alla utfallsmått.

Steg 2: Baserat på resultaten på för-mätningen och råd från den tredje författaren, valde skolan gruppen elever som de ville skulle få träna med *Läsklar*. Om denna grupp var mindre än gruppen som man trodde behövde extra hjälp (t.ex. för att skolan hade begränsade resurser) valdes eleverna med lägst resultat på testen av bokstavskänneteknik och fonologisk medvetenhet (alla eleverna har noll poäng på avkodningstestet).

Steg 3: Inom varje skola matchade forskarna eleverna i par och, när antalet elever var ojämnt, en trippel. Matchningen av eleverna i par/tripplar baserades på de två primära utfallsvariablerna. Men eftersom inga elever kunde avkoda vid för-mätningen, så var det bokstavskännedomstestet som bestämde matchningen i praktiken. Om flera elever hade samma resultat på bokstavskännedom använde vi testet av fonologisk medvetenhet för att bestämma den bästa matchningen. Vi lät sedan slumpen bestämma vilken elev i varje par, eller en eller två elever i varje trippel, som fick träna med *Läsklar* i en första period (dvs ingick i interventionsgruppen). Resten hamnade på väntelistan och fick träna med metoden under en andra period. Den andra författaren utförde randomiseringen till interventions- och kontrollgruppen med hjälp av slumpgeneratorn i Microsoft Excel. Elevernas identiteter var anonymiserade och den enda informationen som var tillgänglig vid randomiseringen var för-mätningens-resultaten.

2.2.4 Är interventions- och kontroll-gruppen lika vid för-mätningen?

Den slumpmässiga indelningen av elever i interventions- och kontrollgrupp gör att vi inte förväntar oss några systematiska skillnader mellan grupperna. Dock kan även korrekta randomiseringsförfaranden ge upphov till obalanser mellan grupperna, särskilt i små urval, vilket i sin tur kan innebära problem för de statistiska testerna av effekterna (t.ex. Roberts & Torgerson, 1999).

Panel A i Tabell 2 visar medelvärden och standardavvikelser för elevegenskaper och för-mätningens resultat. Alla elever har noll poäng på avkodningstestet, det vill säga ingen elev kunde läsa några ord före insatsen började. I genomsnitt kände deltagarna igen fem bokstäver och hade hälften rätt på testet av fonologisk medvetenhet (fem poäng av högst tio). De flesta elever tyckte att det skulle bli roligt att lära sig att läsa och de ville gärna lära sig: medelpoängen på läsglädje och motivationsmått är cirka åtta (av max tio). Fler elever trodde att det skulle bli svårt att lära sig att läsa, medelvärdet på frågan om självtillit är cirka fem (av tio).

Grupperna är i princip identiska när det gäller de två primära utfallsmåtten, avkodning och bokstavskännedom. Ingen av interventions- och kontrollgruppens medelvärde i panel A, Tabell 2 skiljer sig signifikant från varandra ($p > 0,10$). De två största skillnaderna på för-mätningarna, fonologisk medvetenhet och läsglädje, har olika tecken, vilket tyder på att skillnaderna inte är systematiska.

I vår engelska rapport visar vi vidare att variablerna i Tabell 2, panel A, inte heller gemensamt förutsäger väl vem som hamnade i interventionsgruppen och att de förväntade skillnaderna mellan interventions- och kontrollgruppen på efter-mätningen är små när vi inte tar hänsyn till vem som fick interventionen

(Bøg m.fl., 2019). Sammanfattningsvis verkar därmed randomiseringen ha skapat en interventions- och kontrollgrupp som liknar varandra.

Tabell 2. Medelvärde och standardavvikelser för analysurvalet vid för-, efter-, och uppföljningsmätningarna

Panel A: För-mätning								
Variabel	Intervention			Kontroll			ES	p
	M	SA	n	M	SA	n		
Flicka	0.39	0.49	80	0.51	0.50	76	-0.13	0.12
Årskurs 1	0.20	0.40	80	0.20	0.40	76	0.00	0.97
Specifik risk	0.21	0.41	80	0.26	0.44	76	-0.05	0.46
Avkodning	0.00	0.00	80	0.00	0.00	76	0.00	1.00
Bokstavskännedom	5.11	3.71	80	5.11	3.92	76	0.00	0.99
Fon. medvetenhet	4.25	4.17	80	5.17	4.18	76	-0.22	0.17
Självförtroende	4.87	3.88	80	4.62	3.55	76	0.07	0.68
Läsglädje	8.20	3.24	80	7.35	3.45	76	0.25	0.11
Motivation	8.12	3.41	80	8.27	3.10	76	-0.05	0.77

Panel B: Efter-mätning								
Variabel	Intervention			Kontroll			ES	P
	M	SA	n	M	SA	n		
Avkodning	6.63	7.47	80	1.28	3.55	76	0.90	0.00
Bokstavskännedom	17.86	6.29	80	10.58	7.12	76	1.08	0.00
Fon. medvetenhet	9.00	2.08	80	7.25	3.86	76	0.57	0.00
Självförtroende	6.61	3.45	80	4.66	3.11	76	0.59	0.00
Läsglädje	8.41	3.01	80	8.26	3.06	76	0.05	0.75
Motivation	9.14	2.20	80	8.78	2.35	76	0.16	0.32

Panel C: Uppföljnings-mätning								
Variabel	Intervention			Kontroll			ES	p
	M	SA	n	M	SA	n		
Avkodning	11.79	11.21	72	9.91	12.50	69	0.16	0.35
Bokstavskännedom	21.29	5.99	72	20.23	6.22	69	0.17	0.30
Fon. medvetenhet	9.61	1.01	72	9.16	2.27	69	0.26	0.13
Självförtroende	5.26	3.56	72	6.64	3.08	69	-0.41	0.02
Läsglädje	7.70	3.27	72	8.84	2.30	69	-0.40	0.02
Motivation	8.52	2.87	72	9.14	2.14	69	-0.24	0.15

Note: Medelvärde (*M*), standardavvikelse (*SA*), urvalsstorlek (*n*), skillnaden mellan interventions- och kontrollgruppen uttryckt som en effektstorlek (*ES*), och *p*-värdet från ett tvåsidigt *t*-test av lika medelvärden/proportioner. *ES* är beräknat som Hedges' *g* för alla variabler utom andelen flickor, andelen i årskurs 1 och andelen med specifik risk (se avsnitt 2.5.1 för en beskrivning av Hedges' *g*). *ES* för de senare variablerna uttrycks som skillnaden i andelar.

2.3 Kontrollgruppens undervisning

Eleverna i kontrollgruppen fick skolornas normala undervisning medan de väntade på att få träna med *Läsklar*. I förskoleklass var undervisningen i de flesta skolor baserad på någon version av Bornholmsmodellen (t.ex. Lundberg m.fl., 1988; Häggström, 2007). Bornholmsmodellen använder strukturerade lekar och övningar som syftar till att hjälpa elever att upptäcka språkets fonologiska struktur. Lekarna inkluderar att lyssna på verbala och icke-verbala ljud, rimlek och rimmade berättelser, segmentering av meningar i ord och undersökning av hur långa orden är, handklappning för att markera stavelser, igenkänning av fonem i början och inom ord och prosodiska spel (se t.ex. Lundberg m.fl., 1988). Bornholmsmodellen genomfördes alltid i helklass.

Undervisningen i årskurs 1 baserades på den nationella svenska läroplanen (Skolverket, 2018b). Läroplanen betonar bland annat alfabetisk kunskap, kopplingen mellan bokstäver och ljud och handstil. Särskilt stöd av samma typ som *Läsklar* används sällan i årskurs 1 i svenska skolor. Endast 1,6 procent av elever i årskurs 1 har en individuell utbildningsplan, som är en förutsättning för att ta emot särskilt stöd, och 0,5 procent får en-till-en undervisning (Skolverket, 2018c).

Förskoleklass omfattas inte av den nationella statistiken, men ingen deltagande skola rapporterade att de normalt använder några liknande metoder som riktar sig specifikt mot elever med risk att hamna i lässvårigheter (innan de började använda *Läsklar*). Vissa skolor använder en-till-en eller smågruppsundervisning för elever med lässvårigheter i den senare delen av första klass. Alla årskurs 1 elever i interventionsgruppen tränade dock med metoden under första halvåret av årskurs 1. Därför verkar det osannolikt att studien skulle ha påverkat den normala undervisningen negativt för kontrollgruppen, i den meningen att kontrollgruppseleverna skulle ha fått särskilt stöd under samma period, om vi inte hade genomfört studien.

2.4 Implementering

Vi övervakade inte hur skolorna genomförde insatsen för att påverka så lite som möjligt och komma nära hur skolorna använder *Läsklar* utanför studien. De medverkande lärarna kunde kontakta den tredje författaren om de hade frågor, men det skulle ha varit möjligt också om skolorna hade genomfört programmet utanför studien (3–5 skolor använde denna möjlighet). Vi intervjuade ett urval av lärarna efter uppföljningsmätningen (fem skolor saknas i intervjuunderlaget). De flesta skolor som inte hade använt *Läsklar* tidigare var inte vana vid att genomföra liknande metoder och organiseringen sågs till en början som proble-

matisk, eftersom det inte fanns några etablerade rutiner. Samtliga skolor lyckades emellertid lösa dessa problem, vilket framgår av att det genomsnittliga antalet genomförda sessioner ligger nära det avsedda antalet. Mycket få elever fick ett antal sessioner som var väsentligt under de avsedda 30–35 (t.ex. är det bara 3 elever, som fick mindre än 25 sessioner). Båda de deltagande lärarna på en skola bytte jobb efter att interventionsgruppen hade tränat med metoden. Skolan kunde därför inte ge metoden till eleverna på väntelistan och vi genomförde inga uppföljningstest. En annan skola hade fler elever som man ansåg behövde metoden än vad de hade resurser till. Vi kom överens med skolan om att låta den tredje författaren träna åtta elever. Dessa elever ingår i våra resultat, men vi visar i vår engelska rapport att resultaten är närmast identiska om vi istället utesluter dem (Bøg m.fl., 2019).

2.5 Analysplan

Analysen följer vår registrerade analysplan, som finns tillgänglig i American Economic Association's RCT Register (www.socialscienceregistry.org) med registreringsnummer AEARCTR-0002750. På grund av den sekventiella rekryteringen av skolor registrerades planen efter att försöket hade börjat men innan forskarna som utförde den statistiska analysen (den första och andra författaren) hade sett data och innan alla skolor avslutat träningen med sina elever.

I vår engelska rapport presenterar vi ett antal känslighets- och heterogenitetsanalyser (Bøg m.fl., 2019). Vi kommenterar kort några av dessa nedan men eftersom våra huvudresultat inte är märkbart känsliga för hur vi specificerar den statistiska analysen och evidensen för att effekterna skiljer sig åt mellan olika grupper av elever och skolor är svag, så fokuserar vi på huvudresultaten i den här rapporten.

2.5.1 Analys av de kortsiktiga effekterna

Vi börjar analysen med att visa de fullständiga för- och efter-mätningfördelningarna av samtliga utfallsmått. För att skatta effekterna av interventionen, använder vi oss av linjära regressioner där våra utfallsmått är de beroende variablerna. Vi inkluderar en indikator för att ha blivit randomiserad till interventionsgruppen, en indikator för det par/den trippel eleven blev matchad till, samt för-mätningresultat på alla utfallsmåtten som kontrollvariabler (förutom avkodningstestet eftersom alla elever har noll poäng vid för-mätningen). Våra resultat är inte känsliga för hur vi specificerar denna regression, vi får till exempel väldigt lika resultat om vi bara jämför medelvärdena i interventions- och kontrollgruppen (se Bøg m.fl., 2019).

Vi använder också ett mått på hur stora effekterna är, som kallas Hedges' g (Hedges, 1981). Måttet uttrycker effekten i termer av standardavvikelse på varje utfallsmått (korrigerat för urvalsstorleken). Eftersom effekter från alla typer av test kan omvandlas till standardavvikelse så är g ett bättre mått än de "råa" testpoängen, när vi vill jämföra effekter mellan studier.

2.5.2 Kostnader och kostnadseffektivitet

Syftet med utvärderingen av kostnader och kostnadseffektivitet är att få en uppskattning av hur mycket extra det kostar en skola att genomföra en liknande intervention och få samma effekter på elevernas läskunnighet.⁵ Vi följer metoderna i Hollands m.fl. (2013, 2016), Levin och Belfield (2015), Jacob m.fl. (2016) och uttrycker kostnadseffektiviteten som kostnaderna per enhets ökning av vårt mått på effekternas storlek (dvs g).

Vi beräknar kostnaderna för en typisk skola i vårt urval och använder tre överordnade kategorier: personalkostnader, lokalkostnader och kostnader för material och utrustning. Vi inkluderar tre typer av personalkostnader. Den första mäter tiden det tog för lärarna att lära sig hur man genomför metoden. Den andra är den tid som användes för att förbereda och genomföra sessionerna. Den tredje kategorin avser tiden det tog att göra för-mätningen och urvalet av eleverna. Eftersom efter-mätningen var en del av studien men inte är en del av metoden räknar vi inte med kostnaderna för dessa test. Sådana mätningar är antingen något som skolan redan gör eller inte vill göra, oavsett så är den extra kostnaden noll.

Vi samlade in information om personalkostnader genom att intervjua lärare och vår testledare, samt med en kort enkät till lärarna. Vi konverterar lärartimmar till SEK/USD/Euro genom att använda information om den genomsnittliga lönen för lärare, specialpedagoger och speciallärare i Sverige med liknande typ av erfarenhet som de som deltog, samt information om köpkraftsparitetsjusterade valutakurser för 2017 (OECD, 2018).⁶ Eftersom växelkurserna kan fluktuera mellan år konverterar vi också alla kostnader till timmar med hjälp av en viktad kostnad per timme, där kostnaden speglar andelarna av de tre typerna av lärare

⁵ Dvs, vi räknar inte med kostnader som skolan alltid har (t ex kostnader för att anställa lärare), utan bara sådana där det kostar något utöver det vanliga eller kostnader för resurser som skulle kunna använts till annat (t ex kunde den tid som lärarna använde för att lära sig och genomföra metoden använts till annan undervisning).

⁶ De genomsnittliga lönerna är från november 2017 och gäller erfarna lärare, specialpedagoger, och speciallärare i de tidiga åren av grundskolan (kommunikation med Lärarförbundet). Vi beräknar timkostnaden på följande sätt: $((11 \times \text{månadslön} + 1 \text{ månads semesterlön}) \times 1.3142 [\text{arbetsgivaravgifter}] \times 1.05 [\text{arbetsgivares försäkringskostnader}]) / 1760 [\text{årligt antal arbetstimmar}]$ (se t ex Bolagsverket, 2018). Beräkningarna ger timkostnader på 328 SEK (lärare), 375 SEK (specialpedagog), and 369 SEK (speciallärare).

som deltog i vår studie. I de få fall där vi saknar information, beräknar vi antingen värden med medelvärdet för skolan eller för alla skolor (t.ex. om det inte var möjligt att få information från en hel skola).

Vi har inte tillgång till uppskattningar av alternativkostnaderna för att använda skollokaler. Det är oklart om lokalerna skulle ha använts för något annat, men vi använder en uppskattning på 6 USD per elev (dvs ca 51 SEK). Denna kostnad motsvarar ett viktat genomsnitt av lokalkostnaden per elev i de två mest lika insatserna i Hollands m.fl. (2013; 3 USD per elev i *Stepping Stones*, som varar 5 veckor och 11 USD per elev i *Sound Partners*, som varar 18 veckor).

Material och utrustning inkluderar kostnader för *Läsklar*-materialet och utbildningen, kostnaderna för att kopiera bilderna som används i den andra typen av session och kostnaderna för testerna. Två skolor hade redan använt *Läsklar* tidigare och tre andra skolor använde metoden i en andra omgång för en ny årgång elever. Vissa kostnader, som utbildning av lärare och material, var därför inte relevanta för dem, så vi använder de genomsnittliga kostnaderna i andra skolor för de två skolorna i dessa kategorier.

Vi beräknar den totala kostnaden per elev i interventionsgruppen som det viktade genomsnittet över skolor, med vikter som motsvarar skolornas andelar av eleverna i interventionsgruppen. De totala kostnaderna bör därför tolkas som den genomsnittliga kostnaden per elev i interventionsgruppen för att genomföra metoden första gången en genomsnittlig skola i vårt urval använder den.

2.5.3 Resultat vid uppföljningsmätningen

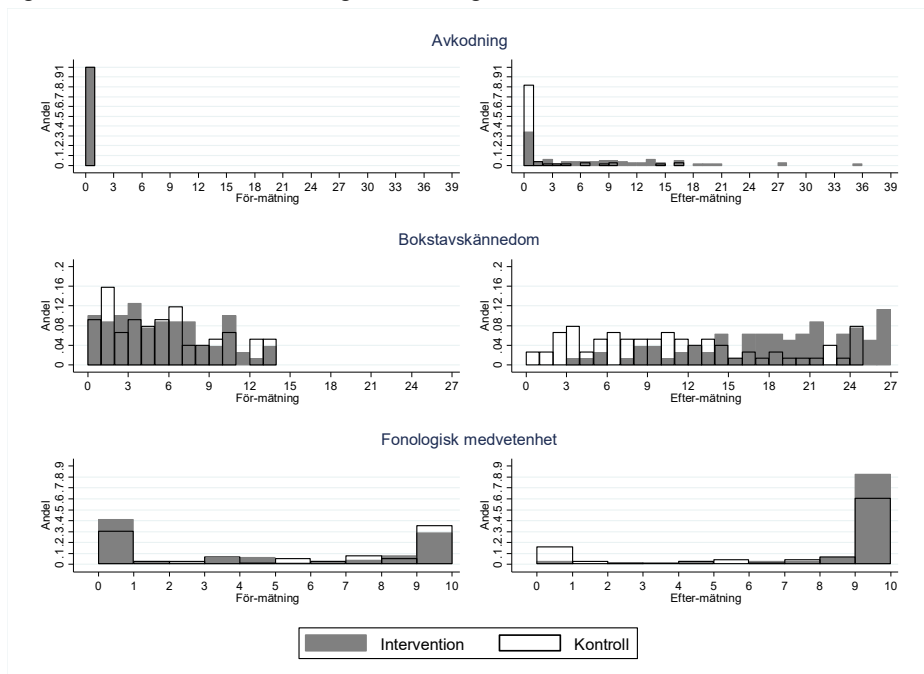
Vi visar förändringarna i utfallsmåtten mellan efter- och uppföljningsmätningen och diskuterar skillnader mellan interventions- och kontrollgruppen. I vår engelska rapport visar vi också resultaten från en regressionsanalys, där vi försöker undersöka om det var bättre att få insatsen i första eller andra omgången (Bøg m.fl., 2019). Av olika anledningar som vi diskuterar ingående i den engelska rapporten är dessa resultat svåra att tolka rakt av som ett test av om det var bättre att ingå i interventionsgruppen eller i att hamna på väntelistan. Av denna anledning och för att indikationerna på skillnader mellan grupperna är relativt svaga, så utelämnar vi dessa resultat i denna rapport.

3 Resultat

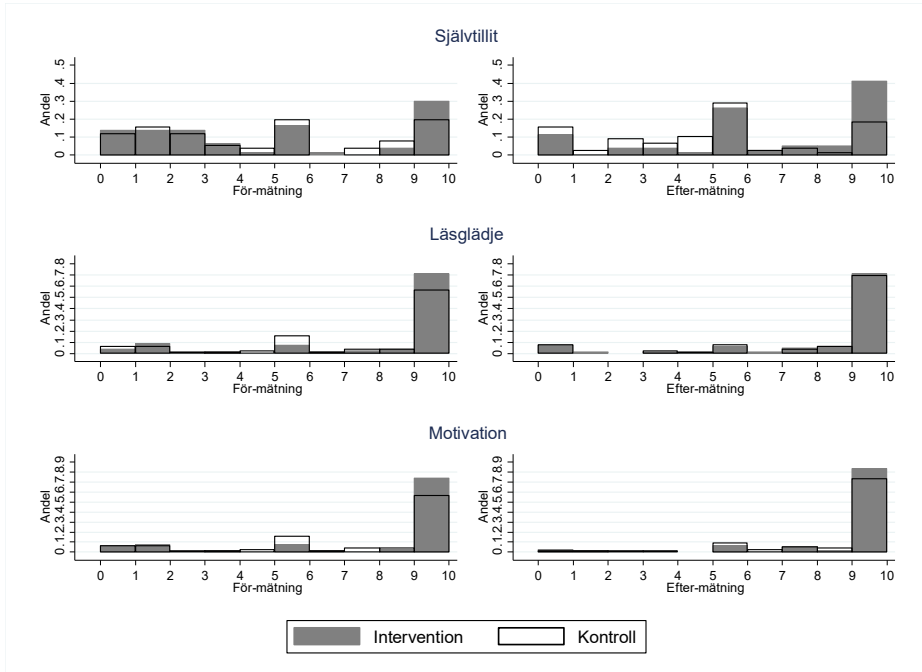
Figur 1 och Figur 2 visar fördelningarna av utfallsmåtten vid för- och eftermätningen. Poängen för varje mått visas på x-axeln och andelen elever i ett visst intervall på y-axeln. De grå staplarna representerar interventionsgruppen och de vita staplarna kontrollgruppen. Alla fördelningar är likartade för de två grupperna vid för-mätningen. Figurerna visar vidare att interventionsgruppen förbättrar sina resultat mer än kontrollgruppen över stora delar av fördelningarna från för- till efter-mätningen på de tre läsrelaterade testen och självtillits-måttet. Det finns små skillnader både över tid och mellan grupperna på måtten av läsglädje och motivation. De flesta eleverna har höga poäng redan vid för-mätningen och poängen ökar något vid efter-mätningen. Det innebär att det finns en klar risk för takeffekter på dessa två test, dvs att vi inte upptäcker förbättringar i de underliggande attityderna för att eleverna når maxpoängen.

Faktum är att en rimligt stor andel av eleverna når maxpoängen även på bokstavskänedom, fonologisk medvetenhet och självtillit. De flesta kontrollgrupps-elever har fortfarande noll poäng på avkodningstestet vid efter-mätningen. Andelen elever som har noll poäng är mycket mindre i interventionsgruppen; omkring 65 procent kan avkoda minst ett ord.

Figur 1 För- och efter-mätningfördelningarna, läsrelaterade tester



Figur 2 För- och efter-mätningfördelningarna, attitydmått



Tabell 3 visar skattningarna av de kortsiktiga effekterna av *Läsklar*. Varje kolumn motsvarar ett av de sex utfallsmåtten. Koefficienterna på variabeln som indikerar om en elev var i interventionsgruppen eller inte visar effekterna uttryckt som testpoäng (robusta standardfel visas inom parentes under varje koefficient).

Effekterna av interventionen är positiva och statistiskt signifikanta i kolumnerna (1) och (2). Elever som fått träna med *Läsklar* kan i genomsnitt avkoda ungefär sex ord mer och känner till sju bokstäver mer än kontrollgruppen ($p < 0,001$ för båda måtten). Effekterna uttryckta som Hedges' g visas i tabellens nedre del och är $g = 1,07$ och $g = 1,03$. (Vi diskuterar storleken på effekterna i avsnitt 4.)

Effekterna är också positiva och signifikanta för fonologisk medvetenhet ($g = 0,56$, $p < 0,001$) och självttillit ($g = 0,57$, $p < 0,001$). Effekterna är mindre och inte statistiskt signifikanta när det gäller läsglädje ($g = -0,08$, $p = 0,623$) och motivation ($g = 0,11$, $p = 0,374$).

Vi visar också koefficienterna på för-mätningvariablerna, vilka ingår som kontrollvariabler i regressionerna. Det mesta av variationen fångas dock av indikatorerna för par/tripplar och få för-mätningar korrelerar starkt med utfallen. Ett

undantag är testet av fonologisk medvetenhet, vilket är starkt korrelerat med de tre läsrelaterade testerna.

Tabell 3 Kortsiktiga effekter

Variabler	Av-kodning (1)	Bokstavs-kännedom (2)	Fonologisk medvetenhet (3)	Självtillit (4)	Läsglädje (5)	Motivation (6)
Intervention	6.333*** (1.123)	6.911*** (0.835)	1.731*** (0.454)	1.767*** (0.542)	-0.240 (0.486)	0.342 (0.382)
<i>För-mättningsresultat</i>						
Bokstavs-kännedom	-0.619 (0.575)	0.234 (0.479)	0.282 (0.211)	-0.694** (0.325)	0.0256 (0.254)	-0.00180 (0.169)
Fonologisk medvetenhet	0.470*** (0.148)	0.354** (0.157)	0.333*** (0.0777)	0.158 (0.117)	-0.0670 (0.111)	0.0168 (0.0612)
Självtillit	-0.265 (0.167)	-0.181 (0.149)	-0.0596 (0.0891)	-0.114 (0.0992)	-0.0977 (0.0938)	-0.0468 (0.0566)
Läsglädje	-0.327 (0.298)	0.403* (0.227)	0.124 (0.117)	0.135 (0.123)	0.182 (0.117)	0.0438 (0.0805)
Motivation	0.662* (0.343)	-0.0423 (0.158)	-0.0791 (0.113)	0.394** (0.179)	0.250* (0.134)	0.147 (0.0894)
Hedges' g	1.07	1.03	0.56	0.57	-0.08	0.11
Antal obs.	156	156	156	156	156	156

Not: Tabellen visar koefficienter och robusta standardfel (i parenteser) från linjära regressioner (beskrivna i avsnitt 2.5.1.). *** indikerar statistisk signifikans på 1-procentsnivån ($p < 0.01$), ** på 5-procentsnivån ($p < 0.05$) och * på 10-procentsnivån ($p < 0.1$).

3.1 Kostnader och kostnadseffektivitet

Tabell 4 visar kostnaderna för att genomföra metoden per elev i interventionsgruppen. Det tog i genomsnitt 10,7 timmar att lära sig metoden, förbereda och träna med eleverna. Som nämnts använder vi en uppskattning av kostnader för skollokaler på 6 USD per elev (ca 51 SEK). Material och utrustning kostar 2 202 SEK per elev, vilket inkluderar kostnaden för materialet, utbildningen, kopior och kostnaderna för testerna. Vi har använt det rådande priset som skolorna skulle betala för att köpa *Läsklar* utanför studien, en uppskattning av kopieringskostnaden och det pris som betalades av studien för testmaterialet per

skola. Totalkostnaden uppgår till 6 040 SEK/703 USD/ 627 Euro, eller 17 lärartimmar per elev i interventionsgruppen.

Tabell 4 Kostnader per elev i interventionsgruppen

Kostnader	SEK	USD	Euro	Timmar	%
<i>Personalkostnader</i>	3 786,5	442,9	392,9	10,7	63%
Utbildning av lärare	329,7	38,6	34,2	0,9	5%
Implementering	3 456,8	404,4	358,7	9,8	57%
Urval av elever	116,9	13,7	12,1	0,3	2%
<i>Skollokaler</i>	51,3	6,0	5,3	0,1	1%
<i>Material och utrustning</i>	2 202,2	254,3	228,5	6,2	36%
Material and kurs	1 935,8	226,4	200,8	5,5	32%
Kopior	14,0	1,6	1,5	0,0	0%
Testmaterial	252,4	26,2	26,2	0,7	4%
Totalsumma	6 404,0	703,2	626,7	17,1	100%

Not: Kostnader beräknade enligt beskrivning i avsnitt 2.5.2. Vi har använt följande växelkurser: USD/SEK = 8,549, Euro/SEK = 9,638 (OECD, 2018).

3.2 Resultat vid uppföljningsmätningen

Medelvärdena i panel C i Tabell 2 visar att interventionsgruppen har förbättrat sina poäng ytterligare på de tre läsrelaterade testerna mellan efter- och uppföljningsmätningen. Detta är ett viktigt resultat, eftersom det indikerar att förbättringen av elevernas läsfärdigheter inte var övergående. Attitydmåtten har minskat något jämfört med eftermätningen, men de är fortfarande högre än vid förmätningen för självförtroende och motivation. Efter att de tränat med *Läsklar* har kontrollgruppen förbättrat alla resultat jämfört med eftermätningen. De ligger lite lägre än interventionsgruppen på de läsrelaterade testerna, men har högre poäng på attitydmåtten.

4 Diskussion

I det här avsnittet diskuterar vi storleken på effekterna och jämför kostnadseffektiviteten hos *Läsklar* med andra liknande metoder, för vilka kostnaderna har beräknats med hjälp av liknande metoder som vi använder. Vi diskuterar också hur studiedesignen kan ha påverkat storleken på effekterna samt vår studies begränsningar.

4.1 Perspektiv på effekter och kostnadseffektivitet

För att undersöka hur stora effekterna av *Läsklar* är jämför vi dem med: i) uttalanden från normgivande organisationer; ii) skillnader mellan elevgrupper; iii) hur mycket elever vanligtvis lär sig på ett år; och iii) effekter av andra insatser. Vi begränsar diskussionen till de tre läsrelaterade testerna, eftersom grupp i) till iii) inte finns tillgängliga för attitydmåten och det är oklart hur jämförbara våra mått är med de få andra mått som använts i den tidigare litteraturen.

What Works Clearinghouse i USA menar att effekter över 0,25 standardavvikelse (dvs $g > 0,25$) bör betraktas som ”substantially important” (What Works Clearinghouse, 2014, sida 23). Lipsey m.fl. (2012) visar att amerikanska elever i genomsnitt förbättrar sig 1,52 standardavvikelse mellan förskoleklass och årskurs 1 och 0,97 standardavvikelse mellan årskurs 1 och 2 på standardiserade test i läsning. Skillnaden mellan elever med låg och hög socioekonomisk status är ca 0,7–0,8 standardavvikelse i årskurs 4 (informationen är inte tillgänglig för tidigare årskurser och vi känner inte till liknande statistik för Sverige). Effekterna av *Läsklar* är ca 1 standardavvikelse för våra viktigaste utfallsmått avkodning och bokstavskänning och något större än 0,5 standardavvikelse på fonologisk medvetenhet. Effekterna framstår därmed som stora jämfört med dessa tre riktmärken.

Innan vi jämför effekterna och kostnadseffektiviteten för *Läsklar* med andra insatser vill vi uppmärksamma några viktiga begränsningar av den typen av jämförelser. Det finns ofta skillnader mellan insatsers varaktighet, intensitet, målgrupp, vilken undervisning kontrollgruppen får och mellan de test som används för att undersöka effekterna. Det är välkänt att många av dessa faktorer påverkar effekternas storlek (t.ex. Cheung & Slavin, 2016), men svårare att säga exakt med hur mycket när man jämför två individuella studier (vi diskuterar denna fråga vidare i nästa avsnitt). Jämförelseproblemen är värre i analyser av kostnadseffektiviteten, som påverkas både av skillnader i effekter och kostnader samt av skillnader i priser/lön mellan länder och regioner.

Vi har beräknat vårt mått på effekternas storlek med de metoder som används i Dietrichson m.fl. (2017), som är en översiktsartikel över interventioner för elever från familjer med låg socioekonomisk status. Målgruppen i de ingående interventionerna är relativt lik vår målgrupp. Våra två primära utfallsmått är standardiserade tester, som är de enda som ingår i Dietrichson m. fl. (2017). Av dessa anledningar tror vi att denna översikt är den bästa jämförelsen. Den genomsnittliga effekten för alla läsinterventioner i översikten är $g = 0,09$. Den genomsnittliga effekten för alla tutoring-interventioner i läsning är $g = 0,30$. Denna effekt ligger i sin tur nära genomsnittet för de sex studier i översikten som studerade tutoring-interventioner i läsning och där åtminstone en delmängd av deltagarna gick i förskoleklass ($g = 0,31$).⁷ Således är våra effekter också stora i jämförelse med både de genomsnittliga effekterna av andra skolinterventioner och liknande tutoring-interventioner.

För att minimera skillnaderna mellan studierna, jämför vi kostnadseffektiviteten av *Läsklar* med några av de få interventioner som innehåller tutoring, riktar sig mot liknande riskgrupper i liknande årskurser och där det finns beräkningar av kostnadseffektiviteten tillgängliga. Vi fokuserar jämförelsen på testen av avkodning och bokstavskänedom eftersom de är våra två primära utfallsmått, de är standardiserade och de har liknande motsvarigheter bland standardiserade tester som använts i tidigare interventioner från USA (båda är tester av "alphabetic" i terminologin som Hollands m.fl., 2013, använder).

I genomsnitt är kostnadseffektiviteten för *Läsklar* 671 USD per standardavvikelses ökning av testresultaten för avkodning och bokstavskänedom. Hollands m.fl. (2013) beräknar kostnadseffektiviteten av tre jämförbara program: *Stepping Stones*, en 5-veckors insats, där effekterna undersöktes i förskoleklass i Nelson m.fl. (2005a, b); *Sound Partners*, en 18-veckors insats utvärderad i förskoleklass i Vadasy and Sanders (2008); och *Reading Recovery*, en 20-veckors insats i årskurs 1, som utvärderats i Schwartz (2005). *Stepping Stones* är något mer kostnadseffektivt än *Läsklar* med 570 USD per standardavvikelses ökning i effekter på test av "alphabetic". Både *Sound Partners* och *Reading Recovery* är mindre kostnadseffektiva (2 093 och 1 480 USD per standardavvikelses ökning av effekterna på test av "alphabetic").

Två stycken årskurs 3-program (*Corrective Reading* och *Wilson Reading*) som också undersöktes i Hollands m.fl. (2013, 2016) är jämförelsevis betydligt mindre kostnadseffektiva, men är också riktade mot en bredare uppsättning läsförmågor. Jacob m.fl. (2016) utvärderar en version av *Reading Partners*, som

⁷ Denna genomsnittseffekt är baserad på följande studier: Burns m.fl. (2003), Vadasy & Sanders (2008, 2010), Amendum m.fl. (2011), Nielsen m.fl. (2012) och Apel & Diehm (2014). Några av dessa studier inkluderar också elever i högre årskurser än förskoleklass.

använder volontärer som tutors och riktar sig mot riskgruppselever i årskurs 2–5. Volontärerna minskar kostnaderna för skolorna, men kostnaden per standardavvikelses ökning av effekterna är fortfarande mycket större än för *Läsklar*, 6 455 USD för de test som är mest lika våra (orsaken är främst att effekterna är små). Både när det gäller effekter och när det gäller kostnadseffektivitet står sig därmed *Läsklar* väl i jämförelse med andra liknande metoder.

4.2 Studiedesign och effekter

Effekternas storlek kan påverkas av hur studien är utformad (t.ex. Cheung & Slavin, 2016). Nedan diskuterar vi de fyra delar av vår studiedesign som vi tror är de viktigaste i detta avseende och jämför med de sex mest liknande tutoring-studierna inkluderade i Dietrichson m.fl. (2017; ”jämförelsestudierna” i texten nedanför, se fotnot 6 för referenser).

Kontrollgruppen fick i vår studie skolornas vanliga undervisning medan de väntade på att få träna med metoden. Den vanliga undervisningen innefattade inte särskilt stöd eller tutoring. Vidare, att kontrollgruppen (och deras föräldrar och lärare) visste att de skulle få träna med metoden kan ha minskat risken för att kontrollgruppen fick mer hjälp utanför skolan jämfört med studier där kontrollgruppen inte alls får någon insats. Det vill säga, vår design borde ge större effekter än i studier där kontrollgruppen får åtminstone lite extra hjälp. Av jämförelsestudierna liknar vår design den i Vadasy and Sanders (2008, 2010) och Apel och Diehm (2014), medan det är oklart vad kontrollgruppen får i de andra tre. Kontrollgruppen i jämförelsestudierna förbättrar sig i genomsnitt 0,68 standardavvikelse mellan för- och efter-mätningarna. I jämförelse förbättrar sig vår kontrollgrupp i genomsnitt 0,55 standardavvikelse på de två primära utfallsmåtten.

Våra primära utfallsmått är standardiserade tester och de är inte konstruerade specifikt för studien. Innehållsmässigt ligger de dock fortfarande relativt nära innehållet i *Läsklar*, vilket tenderar att ge större effekter (Cheung & Slavin, 2016). Vår kontrollgrupp fick emellertid också undervisning inom dessa områden (se avsnitt 2.3.2). Den genomsnittliga effekten i jämförelsestudierna baseras också på standardiserade tester. För de test i jämförelsestudierna där vi kan jämföra är vår bedömning att över 80 procent är ungefär lika väl anpassade. Vi hittade inget test som inte alls motsvarar innehållet i interventionerna i jämförelsestudierna.

Vi tillhandahöll minimalt med extra stöd till skolorna och vi övervakade genomförandet i liten grad, för att så långt som möjligt efterlikna hur *Läsklar* implementeras när skolor använder metoden utanför studien. Dessa delar förvän-

tar vi ger mindre effekter (se t.ex. Thomas m.fl., 2018, för en intressant diskussion och resultat i linje med denna hypotes). Alla jämförelsestudier övervakar/stöder de deltagande skolorna mer än vi gör eller har anställt egen personal som sköter genomförandet av insatserna.

Vi inkluderar de elever som har störst risk att hamna i lässvårigheter. Vadasy and Sanders (2008) och Nielsen m.fl. (2012) använder urvalsprocedurer som bör resultera i att de inkluderar en liknande riskgrupp, medan de andra jämförelsestudierna antingen utesluter de elever som har störst risk eller inkluderar grupper av elever som har mindre risk att hamna i lässvårigheter. Det är dock inte klart huruvida det är svårare eller lättare att få stora effekter med elever som har större risk att hamna i lässvårigheter.

Sammantaget har vår studiedesign både delar som talar för större och mindre effekter i förhållande till jämförelsestudierna. Kombinationen av insatskomponenter i *Läsklar*, snarare än studiedesignen, skulle därmed kunna vara orsaken till de stora effekterna. Vår design tillåter oss emellertid inte att utvärdera effekterna av enskilda insatskomponenter eller att separera dessa effekter från effekterna av studiedesignen. Våra effekter gäller interventionen som helhet. Multisensoriska metoder, olika gruppstorlekar och olika typer av tutors skulle vara intressanta komponenter att undersöka i en design där insatskomponenter testas mot varandra medan studiedesignen hålls konstant.

4.3 Begränsningar

Vi hade för avsikt att komma så nära som möjligt hur skolorna skulle ha genomfört *Läsklar* utanför studien. För att undersöka effekterna var vi dock tvungna att utföra mätningar, vilket troligtvis skolorna inte skulle gjort på samma sätt om de inte deltog i studien. De deltagande lärarna var vidare medvetna om att de deltog i en studie och eftersom föräldrarna informerades om studien var de också medvetna om att studien pågick. Det kan ha skapat så kallade Hawthorne effekter – till exempel kan lärare ha ansträngt sig mer för att de visste att de deltog i studien – men den typen av problem är mycket svåra att undvika i utbildningsinsatser och vår studie skiljer sig inte från andra i detta avseende.

Vår väntelistadesign ger oss inte möjlighet att utvärdera metodens långsiktiga effekter. Det finns få tidigare studier som har följt deltagare i skolinterventioner över längre perioder än ett år efter utgången av interventionen. Suggate (2016) dokumenterar i en översiktsartikel om läsinterventioner en positiv men minskande effekt över tid: den långsiktiga genomsnittliga effekten är cirka 60 procent av den kortsiktiga (uppföljningstiden var i genomsnitt ca 11 månader efter interventionen avslutats). De två studier med längst uppföljningstid bland de som riktar sig specifikt mot elever med risk för lässvårigheter – Blachman m.fl. (2014)

och Elbro och Petersen (2004) – finner båda statistiskt signifikanta effekter tio och sju år efter interventionerna. Dessa resultat tyder på att effekterna av riktade insatser kan vara långvariga. Det är i vårt fall också betryggande att interventionsgruppen fortsätter att förbättra sig på alla läsrelaterade test mellan efter- och uppföljningsmätningen och att effekter kring 60 procent av våra kortsiktiga effekter fortfarande skulle vara betydande. Två aspekter av *Läsklar* kan dessutom vara till hjälp för eleverna att komma ihåg vad de lärt sig: användningen av handalfabetet hjälper dem att komma ihåg kopplingen mellan ljud och bokstäver och tidpunkten för interventionen säkerställer att bokstavskänedom och avkodning fortfarande är i fokus i den ordinarie undervisningen.

Våra utfallsmått har några begränsningar. Attitydmåtten är inte psykometriskt validerade. Därför bör resultaten baserade på dessa mått tolkas med försiktighet och det är osäkert vilka underliggande attityder som måtten fångar. Avkodningstestet innehåller ett par ord som överlappar med dem som används under träningen med *Läsklar*. Vi visar dock i vår engelska rapport att dessa ord inte driver våra resultat (Bøg m.fl., 2019). Våra mått kan dessutom ha problem med golv- (avkodningstestet) och takeffekter (alla andra test, men framförallt läsglädje- och motivationstesterna och mer så vid uppföljningsmätningen än vid eftermätningen). Golveffekter kan överskatta effekterna om fler kontrollgrupps elever ligger nära att få en poäng i stället för noll och underskatta effekterna om det istället är fler i interventionsgruppen som är närmare att knäcka läskoden. Omvänt kan takeffekter underskatta effekterna, åtminstone när det gäller bokstavskänedom, fonologisk medvetenhet och självtillit, eftersom fler interventionsgrupps elever når maxpoängen på dessa mått.

Våra kostnadsberäkningar baseras på lärarnas minne, inte på en kostnadsdagbok. Här skiljer vi oss inte från de flesta studierna som vi jämför oss med (se Hollands m.fl., 2013). Vi underskattar (överskattar) programmets kostnadseffektivitet om det exempelvis finns positiva (negativa) kamrateffekter, vilket vår studie inte var avsedd att utvärdera. Vi känner inte till någon tidigare studie med trovärdiga skattningar av kamrateffekter av en liknande insats som *Läsklar*.

5 Slutsatser

Vi undersöker effekterna av en läsinlärningsmetod – *Läsklar* – som riktar sig till de elever med störst risk att hamna i lässvårigheter i förskoleklass och årskurs 1. Vi finner positiva och statistiskt signifikanta effekter på våra två viktigaste utfallsmått, ett standardiserat test av avkodning och ett standardiserat test av bokstavskännedom. Dessa effekter är stora: t.ex. är de ungefär tre gånger så stora som genomsnittseffekterna av liknande interventioner i en översikt om interventioner riktade mot elever från familjer med låg socioekonomisk status (Dietrichson m.fl., 2017). Effekterna är också positiva och statistiskt signifikanta på mått av fonologisk medvetenhet och självtillit. Effekterna är små och inte statistiskt signifikanta på mått av läsglädje och motivation, vilket skulle kunna förklaras med takeffekter. Insatsen kostar ungefär 6 400 SEK, vilket motsvarar ungefär 17 lärartimmar per elev första gången en skola genomför den (och blir billigare vid ytterligare omgångar), vilket tillsammans med de stora effekterna gör att metoden är mer kostnadseffektiv än de flesta andra liknande insatser för vilka kostnadseffektiviteten har utvärderats (samtliga av dessa insatser är från USA).

Vi anser att dessa resultat är mycket lovande för elever som riskerar att hamna i lässvårigheter. Vårt urval av skolor är geografiskt koncentrerat och inkluderar skolor som valt att delta i studien, eventuellt för att de har större andel elever i riskgrupper än den genomsnittliga svenska skolan. Även om vi inte ser några hinder för att genomföra metoden i andra skolor och de flesta skolor har åtminstone några elever som behöver extra hjälp, så är resultaten inte nödvändigtvis generaliserbara. Nya studier som försöker replikera våra resultat i andra kontexter är därför viktiga. Att undersöka den långsiktiga kostnadseffektiviteten och hur viktig varje komponent i insatsen är skulle också vara intressant framtida forskning.

Källor

- Amendum, S. J., Vernon-Feagans, L., & Ginsberg, M. C. (2011). The effectiveness of a technologically facilitated classroom-based early reading intervention: The targeted reading intervention. *Elementary School Journal*, *112*(1), 107–131.
- Apel, K., & Diehm, E. (2014). Morphological awareness intervention with kindergarteners and first and second grade students from low SES homes: A small efficacy study. *Journal of Learning Disabilities*, *47*(1), 65–75.
- Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*, *106*(1), 46–57.
- Bolagsverket (2018). Räkna ut vad en anställd kostar. www.verksamst.se. Hämtad 2018-09-28.
- Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2018). Incidental learning in a multisensory environment across childhood. *Developmental Science*, *21*(2), e12554.
- Burns, M. K., Senesac, B. V., & Symington, T. (2003). The effectiveness of the HOSTS program in improving the reading achievement of children at-risk for reading failure. *Literacy Research and Instruction*, *43*(2), 87–103.
- Bus, A. G., & IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414.
- Butler, R. (1999). Information seeking and achievement motivation in middle childhood and adolescence: The role of conceptions of ability. *Developmental Psychology*, *35*(1), 146.
- Bøg, M., Dietrichson, J., & Isaksson, A. A. (2018). *A multi-sensory literacy program for at-risk students in kindergarten – Promising results from a small-scale Swedish intervention*. Opublicerat manuskript.
- Bøg, M., Dietrichson, J., & Isaksson, A. A. (2019). *A multi-sensory tutoring program for students at-risk of reading difficulties: Evidence from a randomized field experiment* (IFAU Working Paper 2019:7). Uppsala: Institutet för arbetsmarknads- och utbildningspolitisk utvärdering.
- Cheung, A., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292.

- Cook, P. J., K. Dodge, G. Farkas, R.J. Fryer, J. Guryan, J. Ludwig, & L. Steinberg (2014). *The (surprising) efficacy of academic and behavioral intervention with disadvantaged youth: Results from a randomized experiment in Chicago* (NBER Working Paper no. 19862). Hämtad från <http://www.nber.org/papers/w19862>.
- Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2), 31–47.
- Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *Elementary School Journal*, 104(4), 289–305.
- Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A-M. (2017). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. *Review of Educational Research*, 87(2), 243–282.
- Dietrichson, J., Filges, T., Klokke, R. H., Nielsen, B. C. V., Bøg, M., & Højmark, U. J. (2019). *Targeted school-based interventions for improving reading and mathematics for students with or at-risk of academic difficulties in grade 7 to 12: A systematic review*. Opublicerat manuskript.
- Ehri, L., Nunes, S. R., Willows, D. M., Valeska Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel’s meta-analysis. *Reading Research Quarterly*, 36(3), 250–287.
- Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414–448.
- Elbro, C., & Petersen, D. K. (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.
- Elwér, Å., Fridolfsson, I., Samuelsson, S., & Wiklund, C. (2016). *LäSt – Test i läsförståelse, läsning och stavning för åk 1–6*. Hogrefe Psykologiförlaget: Stockholm.
- Fives, A., Kearns, N., Devaney, C., Canavan, J., Russell, D., Lyons, R., ... & O'Brien, A. (2013). A one-to-one programme for at-risk readers delivered by older adult volunteers. *Review of Education*, 1(3), 254–280.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88(1), 3–17.

- Fälth L., Gustafson S., & Svensson, I. (2017). Phonological awareness training with articulation promotes early reading development. *Education*, 137(3), 261–276.
- Good, R. H., & Kaminski, R. A. (red.). (2002). *Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Eugene, OR: Institute for the Development of Educational Achievement.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, 6(2), 107–128.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177.
- Hollands, F. M., Kieffer, M. J., Shand, R., Pan, Y., Cheng, H., & Levin, H. M. (2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. *Journal of Research on Educational Effectiveness*, 9(1), 30–53.
- Hollands, F. M., Pan, Y., Shand, R., Cheng, H., Levin, H. M., Belfield, C. R., Kieffer, M. J., Bowden, A. B., & Hanisch-Cerda, B. (2013). *Improving early literacy: Cost-effectiveness analysis of effective reading programs*. New York: Center for Benefit-Cost Studies of Education, Teachers College, Columbia University.
- Hoover, W. A., & Tunmer, W. E. (2018). The simple view of reading: Three assessments of its adequacy. *Remedial and Special Education*, 39(5), 304–312.
- Häggström, I. (2007). Att förebygga läs- och skrivsvårigheter med språklekar. In Granström, K. (Ed.) *Forskning om lärares arbete i klassrummet, Forskning i fokus, nr 33*, Stockholm: Myndigheten för Skolutveckling.
- Jacob, R., Armstrong, C., Bowden, A. B., & Pan, Y. (2016). Leveraging volunteers: An experimental evaluation of a tutoring program for struggling readers. *Journal of Research on Educational Effectiveness*, 9(sup1), 67–92.
- Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science*, 14(2), 205–213.
- Johansson, M-G. (2009). *LäsEttan*. Stockholm: Natur & Kultur.
- Levin, H. M., & Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, 8(3), 400–418.

- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., ..., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSE 2013-3000). Washington DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. Hämtad från <http://ies.ed.gov/ncser/>.
- Lovett, M. W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology, 109*(7), 889.
- Lundberg, I., Frost, J. & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263–284.
- Lundberg, I., & Høien, T. (1996). Levels of approaching reading and its difficulties. In B. Ericson & J. Rönnberg (Eds.), *Reading disabilities and its treatment* (pp. 11–33). Linköping, Sweden: Linköping University.
- Machin, S., McNally, S., & Viarengo, M. (2018). Changing how literacy is taught: evidence on synthetic phonics. *American Economic Journal: Economic Policy, 10*(2), 217–241.
- Morgan, P. L., Fuchs, D., Compton, D. L., Córdray, D. S., & Fuchs, L. S. (2008). Does early reading failure decrease children's reading motivation? *Journal of Learning Disabilities, 41*(5), 387–404.
- Muenks, K., & Miele, D. B. (2017). Students' thinking about effort and ability: The role of developmental, contextual, and individual difference factors. *Review of Educational Research, 87*(4), 707–735.
- Nelson, J. R., Benner, G. J., & Gonzales, J. (2005a). An investigation of a prereading intervention on the early literacy skills of children at risk of emotional disturbance and reading problems. *Journal of Emotional and Behavioral Disorders, 13*(1), 3–12.
- Nelson, J. R., Stage, S. A., Epstein, M. H., & Pierce, C. D. (2005b). Effects of a prereading intervention on the literacy and social skills of children. *Exceptional Children, 72*(1), 29–45.
- Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development, 49*(3), 800–814.
- Nielsen, D. C., & Friesen, L. D. (2012). A study of the effectiveness of a small-group intervention on the vocabulary and narrative development of at-risk kindergarten children. *Reading Psychology, 33*(3), 269–299.

- OECD (2016a). *Skills Matter: Further Results from the Survey of Adult Skills, OECD Skills Studies*. OECD Publishing, Paris. Hämtad från https://www.oecd-ilibrary.org/education/skills-matter_9789264258051-en.
- OECD (2016b). *PISA 2015 results (Volume I): Excellence and equity in education*. Paris: PISA, OECD Publishing. Hämtad från <http://www.oecd.org/publications/pisa-2015-results-volume-i-9789264266490-en.htm>.
- OECD (2018). Exchange rates. <https://data.oecd.org/conversion/exchange-rates.htm>. Hämtad 2018-09-28.
- Poskiparta, E., Niemi, P., Lepola, J., Ahtola, A., Laine, P. (2003). Motivational-emotional vulnerability and difficulties in learning to read and spell. *British Journal of Educational Psychology*, 73(2), 187–206.
- Roberts, C., & Torgerson, D. J. (1999). Understanding controlled trials: baseline imbalance in randomised controlled trials. *BMJ*, 319(7203), 185–185.
- Rogde, K., Melby-Lervåg, M., & Lervåg, A. (2016). Improving the general language skills of second-language learners in kindergarten: A randomized controlled trial. *Journal of Research on Educational Effectiveness*, 9(sup1), 150–170.
- Sadoski, M., & Willson, V. L. (2006). Effects of a theoretically based large-scale reading intervention in a multicultural urban school district. *American Educational Research Journal*, 43(1), 137–154.
- Shams, L., & Seitz, A. (2008). Benefits of multisensory learning. *Trends in Cognitive Science*, 12(11), 411–417.
- Shams, L., Wozny, D. R., Kim, R., & Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology*, 2(264), 1–9.
- Skolverket (2018a). <http://www.skolverket.se/statistik-och-utvardering/statistik-i-databaser>. Hämtad 2018-02-26.
- Skolverket (2018b). <https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/laroplan-lgr11-for-grundskolan-samt-for-forskoleklassen-och-fritidshemmet>. Hämtad 2018-09-13.
- Skolverket (2018c). *Statistik om särskilt stöd i grundskolan*. Hämtad från <https://www.skolverket.se/skolutveckling/statistik/arkiverade-statistiknyheter/statistik/2018-04-27-statistik-om-sarskilt-stod-i-grundskolan>.

- Slavin, R. E., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1–26.
- Snow, C., Burns, S., Griffin, P., & the Committee on the Prevention of Reading Difficulties in Young Children, National Research Council (1998). *Preventing reading difficulties in young children*. Washington DC: The National Academies Press.
- Snowling, M., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, 81(1), 1–23.
- Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.
- Statens beredning för medicinsk utvärdering (2014). *Dyslexi hos barn och ungdomar – tester och insatser. En systematisk litteraturoversikt* (SBU-rapport nr 225). Stockholm: Statens beredning för medicinsk utvärdering (SBU). Hämtad från <http://www.sbu.se/sv/publikationer/SBU-utvarderar/dyslexi-hos-barn-och-ungdomar---tester-och-insatser/>.
- Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, 49(1), 77–96.
- Thomas, J., Cook, T. D., Klein, A., Starkey, P., & DeFlorio, L. (2018). The sequential scale-up of an evidence-based intervention: A case study. *Evaluation Review*, forthcoming.
- Tideman, E., Vinneljung, B., Hinze, K., & Isaksson, A. A. (2011). Improving foster children's school achievement: Promising results from a Swedish intensive study. *Adoption & Fostering*, 39(1), 44–56.
- Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-ed.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53, 1–25.
- Vadasy, P. F., & Sanders, E. A. (2008). Code-oriented instruction for kindergarten students at risk for reading difficulties: A replication and comparison of instructional groupings. *Reading and Writing: An Interdisciplinary Journal*, 21(9), 929–963.
- Vadasy, P. F., & Sanders, E. A. (2010). Efficacy of supplemental phonics-based instruction for low-skilled kindergartners in the context of language minority status and classroom phonics instruction. *Journal of Educational Psychology*, 102(4), 786–803.

- Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Paraeducator-supplemented instruction in structural analysis with text reading practice for second and third graders at risk for reading problems. *Remedial and Special Education, 27*(6), 365–378.
- Vaughn, S., Roberts, G., Wexler, J., Vaughn, M. G., Fall, A-M., & Schnakenberg, J. B. (2015). High school students with comprehension difficulties: Results of a randomized controlled trial of a two-year reading intervention. *Journal of Learning Disabilities, 48*(5), 546–558.
- What Works Clearinghouse (2014). *Procedures and Standards Handbook Version 3.0*. Retrieved 2017-02-20 from:
https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standard-handbook.pdf.
- Wolff, U. (2011). Effects of a randomised reading intervention study: An application of structural equation modelling. *Dyslexia, 17*(4), 295–311.
- Wolff, U. (2016). Effects of a randomized reading intervention study aimed at 9-year-olds: A 5-year follow-up. *Dyslexia, 22*(2), 85–100.