

Utvärdering av Sofias lässtart

En multisensorisk läsinlärningsmetod för elever i förskole- och första klass

Martin Bøg¹ & Jens Dietrichson²

Augusti 2016

¹ Martin Bøg är fil. dr. i nationalekonomi och fil. mag i socialantropologi. Han är verksam som forskare vid SFI – The Danish National Centre for Social Research. E-post: mbg@sfi.dk.

² Jens Dietrichson är fil. dr. i nationalekonomi och fil. kand. i statsvetenskap. Han är verksam som forskare vid SFI – The Danish National Centre for Social Research. E-post: jsd@sfi.dk.

Sammanfattning

Att lära barn att läsa är en av skolans viktigaste uppgifter. För barn som har svårt att lära sig läsa är tidiga insatser viktiga, både för att kunna undvika längre perioder av upplevt misslyckande i skolan och för att det troligen är mer resurseffektivt att komma till tals med problem tidigt. Denna rapport undersöker effekterna av en läsinlärningsmetod – *Sofias lässtart* – för elever i förskole- och första klass med behov av extra läsinlärningshjälp. Sofias lässtart innehåller tre huvudkomponenter: ett fokus på fonologisk medvetenhet, multisensorisk träning och intensivundervisning i mindre grupper. Träningen pågår 3-4 gånger i veckan under cirka 8 veckor, i totalt 30 sessioner där varje session varar i ungefär 10-15 minuter. Metoden har hittills använts på tre skolor: Dalhemsskolan och Raus Planterings skola i Helsingborg stad, samt Jens Billeskolan i Bjuvs kommun. Mellan 2011-2016 har totalt 95 elever på de tre skolorna fått Sofias lässtart antingen i förskole- eller första klass, i grupp eller individuellt. Eleverna med störst svårigheter har i regel fått interventionen individuellt.

Vi undersöker effekterna av interventionen på tre standardiserade och normerade tester av avkodningsförmåga, bokstavskänedom och fonologisk medvetenhet. På samtliga test ser vi stora förbättringar från testet som gjorts innan interventionen till det som görs efter. För de tester där vi kan jämföra (avkodning och fonologisk medvetenhet), är interventionsgruppens genomsnitt, både i förskole- och första klass, lika högt eller något högre än de normerade genomsnitten. Eftersom skolorna där Sofias lässtart genomförts har lägre andel elever med föräldrar som har eftergymnasial utbildning och högre andel elever med utländsk bakgrund (i två fall mycket högre) än en svensk genomsnittsskola och eftersom eleverna som fått interventionen i många fall är de med allra störst lässvårigheter på dessa skolor, så är detta anmärkningsvärt starka resultat.

För eleverna som får interventionen i förskoleklass jämför vi också utvecklingen över tid med kontrollgrupper. Vi finner stora effekter av den individuella interventionen på avkodning och bokstavskänedom. Som jämförelse är effekterna lika stora eller större än de genomsnittliga effekterna av de mest framgångsrika interventionstyperna från den internationella forskningen om interventioner för liknande elevgrupper. Effekterna är mindre på testet av fonologisk medvetenhet, vilket skulle kunna förklaras av att delar av kontrollgruppens undervisning ligger närmare de områden som detta test testar. Effekterna är också mindre för eleverna som fått interventionen i grupp. Den troligaste förklaringen till att effekterna är mindre för gruppinterventionseleverna är att dessa har mindre svårigheter när interventionen börjar. När vi begränsar undersökningen till gruppen elever med de lägsta resultaten på för-testen gör dessa minst lika stora framsteg i gruppinterventionen som i den individuella varianten.

Vår analys har ett antal begränsningar, som främst bottnar i att urvalet till interventions- och kontrollgrupper inte är slumpmässigt och att urvalsstorleken är relativt liten i den del av analysen där vi kan jämföra med en kontrollgrupp. Detta gör att våra resultat bör tolkas med viss försiktighet. Sammantaget bör dock resultaten betraktas som väldigt lovande. Särskilt i ljuset av att metoden kräver färre resurser än liknande interventioner i den internationella forskningslitteraturen och att vi inte känner till någon tidigare effektutvärdering av en läsinlärningsmetod, specifikt riktad mot förskole- och förstaklasselever med lässvårigheter i svenska skolor.

Bakgrund

Att lära barn att läsa är en av skolans viktigaste uppgifter, om inte den allra viktigaste. Enligt flera mätningar har läsförståelsen i Sverige uppvisat en nedåtgående trend under de senaste decennierna (se t ex Björklund m fl, 2010; SOU, 2012). De potentiella anledningarna till varför trenden pekar nedåt är många och sträcker sig från förändringar i hur svenska skolor arbetar till större samhällsförändringar som Internets ökade betydelse. Oavsett orsak gör den nedåtgående trenden det mycket angeläget att hitta effektiva program och metoder för att hjälpa barn som har svårt att lära sig läsa.

Tidiga insatser är troligtvis att föredra av åtminstone två anledningar: Om problem kan åtgärdas tidigt, eller ännu bättre kan förebyggas, behöver barn inte uppleva längre perioder av skolmisslyckande. Chansen att läsning och, givet läsningens betydelse för i princip samtliga övriga skolämnen, skolarbete förknippas med lust istället för olust bör därför öka ju tidigare läsinlärningssvårigheter kan åtgärdas. Det är också troligt att det är resurseffektivt att ingripa tidigt. Effektiviteten av senare investeringar i elevers humankapital kan vara beroende av den tidigare nivån (Heckman, 2000) och därmed att tidiga insatser genomförs. Eftersom långa perioder av skolmisslyckande rimligtvis är negativt för elevers motivation och självbild, kan senare insatser behöver vara mer omfattande för att lyckas (Cook m fl, 2014). Även om det finns exempel på effektiva interventioner för äldre elever i den internationella forskningslitteraturen, så är dessa exempel klart färre än exemplen på interventioner som visat sig vara effektiva för yngre elever (se t ex Wanzek m fl, 2013; Scammaca m fl, 2015; och Dietrichson m fl, 2016, för översikter).

Det finns dock endast få exempel på läsinlärningsmetoder riktade mot elever med lässvårigheter i svenska skolor, vilka också har utvärderats på ett sätt som gör det möjligt att uttala sig om metodernas effekter. Ett undantag är Wolff (2011, 2016) som beskriver resultat från en randomiserad intervention med fonologisk träning för elever med lässvårigheter i tredje klass. Interventionsgruppen hade signifikant bättre läsförmåga inom ett antal olika områden direkt efter interventionen och även fem år senare när det gällde avkodning. För de allra yngsta eleverna i förskole- och första klass känner vi inte till någon effektutvärdering gjord i svenska skolor under de senaste decennierna.³

Denna rapport undersöker effekterna av en läsinlärningsmetod – *Sofias lässtart* – utvecklad av specialpedagog Anna Aldenius-Isaksson.⁴ Rapporten är gjord på uppdrag av Anna Aldenius-Isaksson och finansierad med medel från Almi innovation. Sofias lässtart riktar sig mot elever i förskoleklass och första klass med behov av extra läsinlärningshjälp. Metoden innehåller tre huvudkomponenter: ett fokus på fonologisk medvetenhet, multisensorisk träning och intensivundervisning i mindre grupper (vi använder det engelska begreppet ”tutoring” för den senare typen av undervisning). Framförallt tutoring och fonologisk medvetenhet har stark

³ Många svenska skolor arbetar med ”Bornholmsmodellen”, som har ett liknande fokus som på fonologisk medvetenhet som Sofias lässtart och har utvärderats med goda resultat (Lundberg, 1988). Bornholmsmodellen genomförs dock i regel i större grupper och är inte speciellt riktad mot barn med läsinlärningssvårigheter, även om effekterna ser ut att vara större för barn i riskgrupper (Hägström, 2007). Samtliga skolor som använt Sofias lässtart arbetar också enligt Bornholmsmodellen, så de effekter vi ser av Sofias lässtart är utöver de som fås av Bornholmsmodellen.

⁴ Metoden har fått sitt namn av den elev som metoden först utvecklades för, Annas dotter Sofia.

uppsättning i tidigare forskning och har visat sig vara framgångsrika byggstenar för interventioner för elever som är i riskzonen för lässvårigheter.

Fonologisk medvetenhet innefattar att kunna urskilja språkets minsta byggstenar, fonemen, som kombineras till större enheter som stavelser och ord, samt förmågan att kunna uppmärksamma dessa större enheter i talet (Häggström, 2007). I alfabetiska system representeras språkljuden med bokstäverna och det råder relativt stark konsensus inom läsforskningen att fonologisk medvetenhet och bokstavskänedom är två av viktigaste faktorerna för barns tidiga läsutveckling (se t ex Snow m fl, 1998; Snowling & Hulme, 2011).

I en stor forskningsöversikt och meta-analys fann Slavin m fl (2011) att tutoring-interventioner inriktade på att öka den fonologiska medvetenheten är den typ av läsinterventioner som har störst effekter för yngre elever med lässvårigheter. Tutoring-interventioner i allmänhet hade även de största effekterna i en ny översikt som undersökte effekter av interventioner för elever från familjer med svag socioekonomisk bakgrund (Dietrichson m fl, 2016).

Multisensoriska metoder syftar till att hjälpa barn med läsinlärningen genom att aktivera flera olika sinnen och därigenom stärka minneskapaciteten. Multisensoriska metoder har länge varit en del av det som anses vara god praxis för att hjälpa barn med dyslexi. Det saknas dock större forskningsöversikter som utvärderar effekterna över många interventioner samtidigt (Snowling & Hulme, 2011). Men det finns många enskilda exempel på framgångsrika interventioner, där multisensorisk träning ingår som en komponent (t ex Sadoski & Willson, 2006; Ehri m fl, 2007).

Resten av rapporten har följande upplägg: I nästa avsnitt beskriver vi innehållet i Sofias lässtart och huvuddragen i den kurs som Anna Aldenius-Isaksson ger för de specialpedagoger/lärare som implementerar metoden. I avsnittet Metod beskriver vi de tre skolor och de elevgrupper som hittills fått träna med metoden, hur vi utvärderar effekterna av metoden och det provmaterial som vi använder som för- och efter-test. I avsnittet Resultat presenterar vi resultaten av utvärderingen och i avsnittet Diskussion jämför vi resultaten med tidigare metoder, diskuterar utvärderingens begränsningar, samt presenterar våra slutsatser.

Sofias lässtart

Sofias lässtart är en multisensorisk läsinlärningsmetod för elever i förskole- och första klass. Metoden riktar sig mot elever som av olika anledningar inte lär sig läsa inom ramen för den ordinarie undervisningen i skolan. Till exempel har barn med följande svårigheter fått insatsen: utvecklingsstörning och andra kognitiva problem, hörselskada, språkstörning, dyslexi, koncentrationssvårigheter och emotionella problem. Metoden har även använts i undervisning av nyanlända barn och ger dessutom träning i handalfabetet.

Metoden utförs av utbildad skolpersonal, vanligen en specialpedagog, och kan ges antingen individuellt eller i mindre grupper, till exempel en specialpedagog och fyra elever. Träningen pågår 3-4 gånger i veckan under cirka 8 veckor, i totalt 30 sessioner där varje session varar i ungefär 10-15 minuter. Som jämförelse kan nämnas att medellängden (och medianlängden) för liknande tutoring-interventioner i Dietrichson m fl (2016) är mer än det dubbla, ca 18-19 veckor. I den

version av metoden som vi utvärderar i denna rapport har specialpedagoger och lärare på 3 skolor i Helsingborgs stad och Bjuvs kommun fått utbildning av Anna Aldenius-Isaksson och sedan genomfört metoden för elever på sina skolor.

Sessionerna i Sofias lässtart följer två steg. I steg 1 används följande material:

- 25 figurer gjorda i cernitlera. Begynnelsebokstäverna i figurerna motsvarar 25 av de svenska språkljuden.⁵
- En låda med 29 fack – ett fack för varje bokstav.
- Ett skrin där figurerna ligger från början.
- Ett inplastat ark med 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 första bokstavsljudet i figurens namn låter. Till exempel ”a” som i apa. När eleven listat ut att apan börjar på a så ska eleven säga a och samtidigt göra ett a med handalfabetet. Handalfabetet hjälper eleven att minnas bokstaven. För att ytterligare hjälpa minnet och stärka kopplingen till bokstaven kan också en spegel användas för att eleven ska kunna se ljudet formas i munnen. När alla tre figurerna har fått flytta in i sitt hus, så är sessionen slut. Denna del tar ungefär 10 minuter. Vid nästa session får eleven först flytta in med de tidigare avklarade figurerna och sedan öka på med tre nya figurer. Det brukar ta ca en månad innan eleven kan de flesta bokstavsljuden.

När eleverna kan ca 15 begynnelsebokstäver till figurerna är det dags att börja med steg 2. I steg 2 används följande material:

- Kopieringsunderlag för 70 bilder på ord med 2-3 bokstäver. Orden är uppdelade i olika svårighetsgrad.
- Sax, klister och tuschpenna.
- A5-pärm (med litet inklistrat handalfabet på insidan) eller vita kort i storlek som visitkort, samt låda till korten.
- Plastfickor i A5-format.
- Ritpapper (tjocka) i A5-storlek.

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 ljud ordet 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 detta. Det gör att eleverna kan öva själva, till exempel hemma, eftersom kontrollen om man har läst rätt finns när man vänder 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. När eleverna är någorlunda säkra på alla bokstavsljud kan man gå över till att läsa vanliga läseböcker eller småböcker om dessa används i den ordinarie läsundervisningen.

⁵ Figurer finns inte för bokstäverna c, x, w och z. Anledningen är att det skulle kunna förvirra barnen, exempelvis vore det logiskt att tänka att citron och zebra skulle börja på s.

Sofias lässtart hjälper därmed elever att lära sig läsa genom att öka deras fonologiska medvetenhet, det vill säga att eleverna förstår att ljudströmmen som kommer när man talar kan delas in i ett antal bokstavsljud och att man kan skilja dessa åt. Metoden använder genom den multisensoriska träningen ett stort antal inlärningskanaler. Tanken är att genom att använda så många sinnen som möjligt så stärks minneskapaciteten. När eleverna hjälper figurerna att flytta in i sina rum får minnet en extra hävstång; dels genom det kinestetiska sinnet (rörelsesinnet), dels genom att minnet får hjälp av att kunna placera bokstavsljudets representant (figuren) i dess egna hus. I början blir det lättare för eleven att se framför sig vilket rum figuren ska vara i än att känna igen bokstavens visuella presentation, men så småningom kopplar eleverna ihop bokstaven med ljudet. Därmed aktiveras också det visuella sinnet på två sätt. Det skapas en liten berättelse kring när figurerna flyttar in i huset, vilket gör att också det episodiska minnet hjälper eleverna. Att eleven också får känna på figuren gör att det taktila sinnet aktiveras. I metoden låter man eleven lyssna ut ljudet utifrån begynnelsebokstaven på figuren och därmed aktiveras det auditiva sinnet. Om eleven har stora svårigheter används också en spegel för att tillsammans med eleven utforska hur det ser ut i munnen när man säger bokstavsljudet. Genom en konsekvent användning av handalfabetet aktiveras också flera sinnen, samtidigt som eleverna får med sig detta redskap som minneshjälp när de går över till den ordinarie textläsningen. När barnen läser kan de också själva ta hjälp av handalfabetet när de försöker minnas någon bokstav som de är osäkra på.

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.

Kurs för pedagoger. Sofias lässtart lärs ut i 5 steg:

1. Teoretisk bakgrund och inövande av handalfabetet. Genomgång i hur materialet för första steget används.
2. Kursledare kommer ut och visar det första steget (med figurerna) tillsammans med elever på den aktuella skolan.
3. Kursdeltagarna visar när de jobbar med det första steget.
4. Kursledare visar det andra steget (pärlen) med elever.
5. Kursdeltagarna visar hur de jobbar med pärlen. Kursdeltagarna erhåller ett utbildningsbevis.

Alla kursdeltagare övar med barn, varpå lärare, barn och kursledare tillsammans reflekterar kring utförandet. Under implementeringen av metoden ges också fortsatt handledning vid behov om det skulle behövas.

Metod

I detta avsnitt beskriver vi först de skolor där Sofias lässtart har använts, de tester vi använder i utvärderingen och sist de metoder vi använder för att undersöka effekterna av interventionen.

Skolor och elever

Sofias lässtart har hittills använts på tre skolor: Dalhemsskolan och Raus Planterings skola i Helsingborg Stad och Jens Billeskolan i Bjuvs kommun. Båda skolorna i Helsingborg hade under läsåret 2015/2016 en hög andel elever med utländsk bakgrund i jämförelse med riksgenomsnitt på 23 procent: Dalhemsskolan hade 78 procent elever med utländsk bakgrund och Raus Plantering hade 82 procent. Jens Billeskolan hade 26 procent elever med utländsk bakgrund. Alla tre skolorna har en lägre andel elever vars föräldrar har eftergymnasial utbildning i jämförelse riksgenomsnittet på 56 procent: andelen är 27 procent på Dalhemsskolan, 25 procent på Raus Plantering och 36 procent på Jens Billeskolan. För läsåret 2014/2015 hade Dalhemsskolan och Jens Billeskolan också en lägre andel elever som uppnådde kunskapskraven i alla ämnen i årskurs 9, 56 procent respektive 68 procent, som kan jämföras med riksgenomsnittet på 77 procent (Skolverket, 2016). Raus Plantering är en F-6 skola så det finns ingen statistik över hur stor andel som uppnått kunskapskraven i årskurs 9.

Sofias lässtart användes först på Dalhemsskolan, med start under 2011. Elever som vid starten av årskurs 1 hade låga resultat på tester av språklig medvetenhet och bokstavskunskap fick arbeta med metoden Sofias lässtart. Under 2014 påbörjades ett försök med att också arbeta med barn i förskoleklass⁶ och under 2015 har en av skolans två grupper i förskoleklass fortsatt detta arbete, riktat mot barn som under vårterminen ännu inte har blivit fonologiskt medvetna. Under 2015 startades även ett försök i förskoleklass på Raus Planterings skola och på Jens Billeskolan.

Projektet i förskoleklass kallas ”Läsberedd i förskoleklass - en multisensorisk lek”. Medan metoden under de första åren endast gavs individuellt av en lärare till en elev och till elever med de största svårigheterna har metoden i förskoleklass getts till elever med mindre läsinlärningssvårigheter och med olika grupp sammansättningar. När metoden använts i smågrupper har dessa i regel bestått av en lärare och fyra elever. För de elever som inte har fått interventionen i förskoleklass har undervisningen på de tre skolorna varit baserad på Bornholmsmodellens strukturerade språklekar (se t ex Häggström, 2007), vilket också är basen för den språkundervisning som interventionsgruppen fått när de inte tränat med Sofias lässtart.

I tabell 1 visar vi antalet elever som lästränat med hjälp av metoden (dvs som fått en intervention) per skola, det år eleverna började förskoleklass, när de fått träningen (förskoleklass eller första klass), samt i vilken form eleverna fått insatsen, individuellt eller i grupp. Sofias lässtart gavs endast i första klass under 2011-2013. De första barnen som fick Sofias lässtart i andra halvan av förskoleklass började skolan hösten 2014. Gruppformen sammanfaller med att förhållandevis fler elever får interventionen. De elever som uppfattades ha störst behov av stöd har i regel fått en individuell intervention, men det finns också elever bland de som fått gruppinterventionen vilka enligt specialpedagogerna som genomfört Sofias lässtart borde fått en individuell intervention om det funnits resurser. Totalt har 54 elever fått interventionen i förskoleklass, 33 av dessa i grupp. Totalt har 41 elever fått interventionen individuellt i första klass. I denna grupp ingår 10 elever som

⁶ Detta försök omfattade fyra elever som började få interventionen i förskoleklass, men vilka sedan får större delen av sessionerna i första klass. Dessa fyra ingår i analysen för första klass-gruppen, eftersom deras första test som görs efter interventionens slut görs i första klass. De ingår inte i kontrollgruppen i den matchade analysen, eftersom de har fått (delar av) interventionen i förskoleklass.

började förskoleklass 2011, den första årgången där Sofias lässtart genomfördes. För dessa elever har vi dock inte tillgång till utvecklingen på jämförbara test och de ingår inte i våra analyser.

Tabell 1. Antal elever per årgång och interventionstyp

	Dalhem					Jens Bille	Raus	Totalt
	2011	2012	2013	2014	2015	2015	2015	
	Åk 1	Åk 1	Åk 1	F/Åk 1	F	F	F	
Antal elever	46	44	39	44	48	26	30	277
Intervention totalt	10	13	13	9	24	15	11	95
Individuellt	10	13	13	4/5	4	6	7	62
Grupp	0	0	0	0	20	9	4	33

Nor: Åren visar vilket år en elevgrupp startade förskoleklass (F). Antal elever anger totalt antal elever i en årskurs och skola vid starten av årskursen då Sofias lässtart genomfördes. Intervention totalt, Individuellt och Grupp anger antalet elever som genomfört Sofias lässtart.

I medeltal har eleverna fått 26 sessioner av Sofias lässtart, men spridningen är relativt stor. De elever som fått minst antal sessioner har fått 10 stycken, medan de som fått flest har fått 39.

Tester

För att testa den språkliga medvetenheten har tre test använts, ett deltest i det normerade testbatteriet LäSt, ett normerat test av hur många bokstäver eleverna känner igen (Bokstavkännedom), samt det normerade testet Hur låter orden. Alla test som vi använder för att utvärdera Sofias lässtart är därmed standardiserade och inget test är specifikt utvecklat för att mäta effekterna av just denna intervention.⁷

Samtliga tre test har använts innan starten av interventionen för att bedöma vilka barn som har störst svårigheter och därmed störst behov av Sofias lässtart. De tre testen har i olika omfattning också genomförts efter att eleverna fått träna med Sofias lässtart. Alla elever har gjort testet i nära anslutning till interventionens slut och i några årgångar/skolor har fler uppföljningstester genomförts. Eleverna har genomfört testen individuellt med en lärare på skolan (t ex specialpedagog eller klasslärare), som därför i regel varit medveten om en elev fått Sofias lässtart eller inte.

LäSt. LäSt är ett standardiserat och normerat testmaterial för kartläggning och diagnostisering av grundläggande färdigheter i läs- och stavningsförmåga. Materialet innehåller tester av avkodningsförmåga, stavning och läsförståelse. Vi använder ett deltest av avkodningsförmåga (Avkodning ord), där elever läser enstaka ord under en och halv minut och poäng ges för varje korrekt läst ord. Normeringen är gjord med grund i tester av 1 043 elever i Skånes, Västergötlands och Östergötlands län. Vid normeringen har skolor valts ut för att elevunderlaget ska vara så

⁷ Jämfört med test utvecklade specifikt för en intervention, ökar standardiserade test i regel chansen att utvärderingen fångar effekter på läsförmågan som är bredare och inte begränsade till precis det material som eleverna går igenom under själva interventionen (Slavin & Madden, 2011). Effektstorlekarna som uppmätts med standardiserade test är i allmänhet också mindre, se t ex Slavin & Madden (2011) och Scammacca m fl (2015).

representativt som möjligt för årskurs 1 till 6 i Sverige, exempelvis när det gäller socioekonomisk status, annat modersmål än svenska och om eleverna är födda i Sverige (Elwér m fl, 2016).

Bokstavskänedom. Vi använder ett diagnostiskt test från testmaterialet LäsEttan som testar hur eleverna klarar att snabbt benämna bokstäverna i alfabetet. Delprovet tar ungefär 1-2 min att genomföra. *LäsEttan* normerades 2008 på ca 700 elever i årskurs 2 (vi kan därför inte jämföra interventionseleverna med eleverna normeringen är gjord på). Maxpoäng är 27 (Johansson, 2009).

Hur låter orden. Detta material testar den fonologiska medvetenheten hos eleverna. Hur låter orden innehåller tre deltest (Vilket ord är längst? Vilka ord låter lika i början? Hur många ord hör du?). Testmaterialet normerades under 2007-2008 genom att samtliga barn (468) i förskoleklass i Eskilstuna kommun genomförde testet, en gång under september 2007 och en gång i maj 2008. Det maximala antalet poäng på testet är 25 poäng (Hägström & Brännström, 2008).

Inget test är som nämnts specifikt utformat för Sofias lässtart, men det finns skillnader i vad som testas och vad metoden innehåller. Framförallt ligger Bornholmsmodellens språklekar nära innehållet i Hur låter orden (en av författarna till detta test har t ex varit med och utvecklat Bornholmsmodellen). Eftersom eleverna lämnar den ordinarie språkundervisningen när de har sessioner med Sofias lässtart är det möjligt att eleverna i kontrollgruppen har fått mer undervisning som ligger nära innehållet i detta test, vilket är en anledning till att vi förväntar oss större effekter på LäSt och Bokstavskänedom. Vissa elever i kontrollgruppen (i årgångarna 2013-2014) har vidare gjort en mer repetition av Hur låter orden i jämförelse med interventionsgruppen (denna gjordes i starten av förskoleklass). Om fler repetitioner gör det lättare att få ett högt resultat är detta också en anledning att förvänta sig lägre resultat på Hur låter orden.

Analys

Syftet med denna rapport är att undersöka effekten av Sofias lässtart. Effekten av en intervention ska här förstås som skillnaden på ett visst test av att få, respektive inte få, interventionen. Svårigheten att skatta effekter grundar sig främst i att samma grupp/elev inte vid samma tidpunkt både kan få och inte få en intervention. Eftersom vi kan observera resultatet för gruppen elever som faktiskt fick interventionen är det primära utvärderingsproblemet att försöka hitta en trovärdig skattning av den kontrafaktiska situationen, det vill säga hur interventionsgruppens testresultat skulle sett ut om eleverna inte hade fått interventionen. Vi angriper detta problem med två ansatser. Vi diskuterar några av problemen med att skatta effekter av interventioner i detta avsnitt, men återkommer också till denna diskussion i det avslutande avsnittet.

Vi tittar först på utvecklingen av testresultaten för de elever som fått arbeta med Sofias lässtart (interventionsgrupperna) från för-test till efter-test(er). Fördelen med denna ansats är att vi, för åtminstone delar av eleverna, kan följa utvecklingen under en längre tidshorisont, då vi har tillgång till data över fler uppföljningstest än bara det som gjorts relativt nära interventionens slut. Nackdelen med att endast undersöka en interventionsgrupps utveckling över tid, utan att jämföra med en kontrollgrupp, är att en sådan ansats blandar samman effekten av interventionen med

elevernas normala progression och riskerar att återspegla regression mot medelvärdet.⁸ Ett sätt att delvis komma till tals med dessa problem är att använda normerade tester och jämföra utvecklingen i interventionsgruppen med den grupp elever som testet är normerat på. I vårt fall avhjälpes en sådan jämförelse förmodligen inte problemen fullt ut. För det första är inte samtliga test normerade både i förskole- och första klass. För det andra riskerar vi fortfarande att blanda samman effekterna av interventionen med effekterna av att gå på en speciell skola.⁹ För det tredje vilar en sådan jämförelse på antagandet att utvecklingen i interventionsgruppen skulle vara det samma som i den elevgrupp som använts i normeringen, om inte interventionsgruppen hade fått interventionen. Eftersom skolorna där Sofias lässtart genomförts skiljer sig ganska kraftigt från en svensk genomsnittsskola, ofta både i termer av socioekonomi och utländsk bakgrund, och eftersom eleverna som fått insatsen valdes ut på grund av att de hade lässvårigheter är det troligt att detta antagande inte håller fullt ut. Om eleverna på de tre skolorna inte skulle ha fått interventionen skulle vi alltså förvänta oss att de hade haft klart lägre resultat än eleverna som testerna är normerade på.¹⁰

En mer trovärdig skattning av interventionens effekt kan därför fås om vi kan hitta en jämförbar kontrollgrupp. Eftersom utvärderingen sker i efterhand har det inte varit möjligt att randomisera elever till interventions- och kontrollgrupper. Vi använder istället en så kallade matchad *difference-in-differences design* (DID, se t ex Heckman m fl, 1998). Först letar vi efter en kontrollgrupp som är lik interventionsgruppen när det gäller resultaten på de tre för-testen. Sedan skattar vi effekten som skillnaden mellan grupperna i medelförändringen mellan för- och efter-testet för varje test.¹¹ Det första steget gör det troligt att vi jämför liknande grupper och det andra steget garanterar att variabler, vars påverkan inte varierar över tid, inte snedvrider våra estimat. Uttryckt på ett annat sätt, så skattar metoden effekten om interventions- och kontrollgruppen skulle ha haft samma genomsnittstrend från för- till efter-test, om interventionsgruppen inte skulle fått interventionen (se t ex Abadie, 2005).

Vår bedömning är att denna metod ger klart bättre möjligheter att skatta effekterna av interventionen. Men metoden begränsar också utvärderingen i flera aspekter: vi kan på grund av datatillgången bara använda de test som görs närmast starten och närmast efter slutet av interventionen och interventionsgruppen begränsas till de elever som börjar förskoleklass 2015 och får interventionen under våren 2016. För att hitta en kontrollgrupp använder vi att eleverna på Dalhemskolan i årgångarna som startade förskoleklass 2013 och 2014 har genomfört de tre testen två gånger under året i förskoleklass (i januari och maj), men inte fått interventionen. Årgångarna 2013 och 2014 utgör en naturlig kontrollgrupp för de elever som startade förskoleklass under hösten 2015 och har fått Sofias lässtart under andra halvan av året i förskoleklass, eftersom många elever förmodligen skulle ha fått interventionen i förskoleklass om denna rutin varit fullt utvecklad vid denna tidpunkt. Fyra elever i årgång 2014 fick Sofias lässtart med start i februari i förskoleklass

⁸ Det vill säga, om de elever som har utvalts till intervention är de som klarat sig sämst på för-testet finns det en risk att eventuella framsteg också avspeglar att dessa elever underpresterade på för-testet.

⁹ Det vill säga, om Dalhemsskolan, Raus Plantering och Jens Billeskolan är i allmänhet bättre eller sämre på att lära barn att läsa fångas även detta upp i jämförelsen och vi fångar inte den ”rena” effekten av att få interventionen.

¹⁰ En omfattande litteratur visar att både låg socioekonomisk status och minoritetsbakgrund i regel samvarierar negativt med skolresultat, se t ex White (1982), Sirin (2005), Skolverket (2009), Björklund & Salvanes (2011) och OECD (2010, 2013).

¹¹ Därav namnet *difference-in-differences*. Såvitt vi vet finns ingen vedertagen svensk översättning för denna metod.

men genomförde huvuddelen av sessionerna i första klass. Dessa elever ingår *inte* i någon kontrollgrupp.¹² Vidare är det bara en grupp av två i Dalhemsskolans årgång 2015, som deltagit i interventionen. Vår ”bruttokontrollgrupp” utgörs därmed av dessa totalt fem grupper (två i årgångarna 2013 och 2014, en i årgången 2015).

Alla elever i bruttokontrollgruppen är dock inte nödvändigtvis jämförbara med alla i interventionsgruppen. När vi undersöker effekterna av att få Sofias lässtart i grupp använder vi *coarsened exact matching* (CEM; Iacus m fl, 2011) för att hitta en kontrollgrupp. Vi använder för-testen på samtliga tre test som matchningsvariabler och inkluderar elever i kontrollgruppen om de har minst en motsvarande elev i interventionsgruppen med ett resultat inom samma intervall på båda testen, och vice versa för interventionsgruppen. Vi väljer intervallgränser med två, delvis motstående, mål i sikte: att behålla interventionsgruppen så intakt som möjligt och att utjämna så mycket av skillnaderna på för-testen som möjligt. Det första målet kräver relativt breda intervall,¹³ men det räcker för att utjämna mycket av skillnaderna. Koefficienterna på för-testresultaten är för samtliga tre test till exempel små och inte signifikanta i en regression där en indikator för interventionsgruppen är utfallsvariabel och de tre testresultaten förklarande variabler. Det vill säga, efter matchningsproceduren är för-testresultaten inte längre predikterande för vilka elever som fått interventionen (se Tabell 3 nedan).

De matchningsvikter som fås av metoden används för att vikta observationerna i den efterföljande DID skattningen. De elever som inte har någon matchande elev i den motsatta gruppen ingår inte i skattningen (de har vikten noll). Vi visar också skattningar där vi inte använder matchningsproceduren utan istället inkluderar alla elever i bruttokontrollgruppen. Eftersom alla elever i gruppen på Dalhemsskolan som ingår i interventionsgruppen har fått interventionen i någon form och relativt många även på Jens Bille och Raus Plantering, så skulle även denna kontrollgrupp kunna utgöra en rimlig skattning av det kontrafaktiska utfallet för interventionsgruppen.

För gruppen som får interventionen individuellt har vi mer information vi kan använda i matchningen. Bedömningen om vem som ska få en intervention görs i regel på grundval av mer information än de test vi har tillgång till. Därför använder vi i matchningen av de som får metoden individuellt att vi vet vilka elever som fick Sofias lässtart individuellt i första klass i årgångarna 2013 och 2014. Dessa elever utgör därmed kontrollgruppen för de som får interventionen individuellt i förskoleklass (vilka tillhör årgång 2015). Vi skattar sedan effekten av interventionen med samma typ av DID modell.

För de elever på Dalhemsskolan som fått insatsen i första klass saknar vi möjligheter att på ett liknande sätt skapa en kontrollgrupp, eftersom det inte finns någon jämförbar grupp som inte fått insatsen och samtidigt gjort ett och samma test två gånger.

¹² Ett problem skulle kunna vara att de elever som istället fick interventionen i första klass kan ha haft en lägre grad av lässvårigheter och därför ha fått insatsen senare. Detta betyder dock att effekten av interventionen förmodligen underskattas, eftersom det är troligt att utvecklingen för dessa elever under förskoleklass är bättre än för elever som har större svårigheter.

¹³ De exakta intervallgränserna är: {10, 20} för LäSt, {8, 16} för Bokstavskännedom och {8, 16} för Hur låter orden.

För att kunna jämföra storleken på effekterna med andra, liknande interventioner använder vi effekttestimaten från vår analys till att räkna fram ett effektstorleksmått, det så kallade *Hedges' g*, enligt följande formel (Lipsey & Wilson, 2011):

$$g = \left(1 - \frac{3}{4N - 9}\right) \times \left(\frac{\bar{X}_1 - \bar{X}_2}{S}\right)$$

I formeln är $\bar{X}_1 - \bar{X}_2$ ett estimat av effekten av interventionen, i vårt fall mätt som skillnaden i genomsnittsförändringen från för- till efter-test mellan interventionsgruppen och kontrollgruppen. S är den gemensamma, ojusterade, standardavvikelsen från efter-testet för interventionsgruppen och kontrollgruppen.¹⁴ Termen i den första parentesen är en justering för små urvalsstorlekar, som görs för att standardavvikelsen har ett nedåtbias i små urval (Hedges, 1981).

Resultat

Detta avsnitt presenterar resultaten av undersökningen i två underavsnitt. Det första beskriver förändringarna inom gruppen som fått Sofias lässtart mellan för- och efter-testen och det andra redovisar resultaten från den matchade DID designen.

Förändring från för- till efter-test

I Tabell 2 visar vi genomsnittet och standardavvikelsen från det test som gjorts närmast före intervention till det som gjorts närmast efter att interventionen avslutats, samt den genomsnittliga förändringen och dess standardavvikelse. Vi har resultat från samtliga tre test och visar resultaten för tre grupper av elever: sammanslaget för alla elever, för de som fått Sofias lässtart individuellt och de som fått interventionen i grupp. Antalet elever som gjort testen skiljer sig åt, främst beroende på att inte alla årskullar gjort alla test och i mindre utsträckning för att enstaka elever inte gjort alla test.

För alla grupper och test ses en relativt stor förbättring från för- till efter-testet. Notera till exempel att förbättringen är större än en standardavvikelse på både LäSt ord och Bokstavskännedom både för eleverna som fått interventionen individuellt och i grupp.

I nästa steg undersöker vi utvecklingen uppdelat på gruppen som fått interventionen i förskoleklass och gruppen som fått interventionen i första klass. Den första gruppen är därmed en blandning av elever som fått Sofias lässtart individuellt och i grupp, medan samtliga i första klass-gruppen har fått interventionen individuellt (samtliga är också från Dalhemsskolan). Figur 1 visar utvecklingen av genomsnittspoängen för förskoleklass-gruppen från för- (repetition 1 i figuren) till efter-testet (repetition 2) på testen Bokstavskännedom, Hur låter orden och LäSt ord. Gruppen förbättrar sig på

¹⁴ Vi använder standardavvikelsen från efter-testet istället för den kanske mer vanliga proceduren att standardisera med för-testet, eftersom både interventions- och kontrollgruppen i vissa skattningar och test har ett medel väldigt nära noll poäng. Standardavvikelsen är därför också mycket liten och som konsekvens blir effektstorleken enormt stor. Vi tror därför att efter-test standardavvikelsen ger ett mer jämförbart effektstorleksmått.

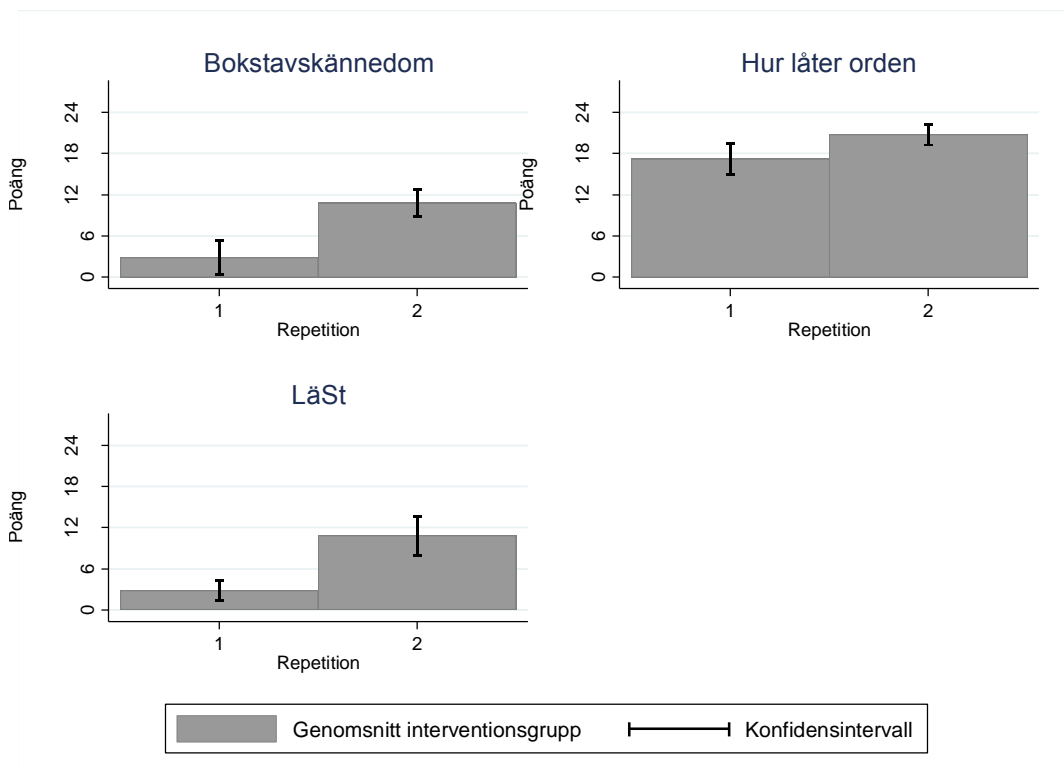
samtliga test, i poäng räknat mest på LäSt ord (7,98), sedan Bokstavskänedom (7,29) och sist Hur låter orden (3,55).

Tabell 2. *Genomsnittlig förändring i interventionsgrupperna från för- till efter-test*

	(1) Antal elever	Före intervention		Efter intervention		Differens	
		(2) Medel	(3) Std.avvikelse	(4) Medel	(5) Std.avvikelse	(6) Medel	(7) Std.avvikelse
<i>Samtliga</i>							
LäSt ord	82	1,76	4,00	15,30	11,60	13,55	11,11
Bokstavskänedom	73	11,96	7,62	21,73	6,49	9,78	5,99
Hur låter orden	47	17,21	7,71	20,77	5,08	3,55	4,26
<i>Individuellt</i>							
LäSt ord	51	0,33	0,89	15,61	12,81	15,27	12,78
Bokstavskänedom	45	9,24	6,48	20,60	7,47	11,36	6,29
Hur låter orden	16	11,19	8,76	16,94	6,35	5,75	5,26
<i>Grupp</i>							
LäSt ord	31	4,10	5,72	14,81	9,46	10,71	6,91
Bokstavskänedom	28	16,29	7,41	23,54	3,98	7,25	4,53
Hur låter orden	31	20,32	4,83	22,74	2,76	2,42	3,18

Not: Kolumn (2) visar medelresultatet på för-testen i en viss grupp och kolumn (3) standardavvikelse för dessa medelresultat. Kolumn (4) visar medelresultatet på efter-testen i en viss grupp och kolumn (5) standardavvikelse för dessa medelresultat. Kolumn (6) är differensen mellan medelresultatet på ett efter-test och ett för-test och i kolumn (7) visas standardavvikelsen av denna differens.

Figur 1. *Utveckling från för- till efter-test i gruppen som fick interventionen i förskoleklass*



I samtliga fall är genomsnittet på efter-testet signifikant större än genomsnittet på för-testet, vilket visas av att den nedre gränsen i det 95-procentiga konfidensintervallet på efter-testet är högre än den övre gränsen på för-testets intervall. Denna ordning gäller också för hur hög genomsnittspoängen var på för-testet, vilket är en tänkbar förklaring till att förändringen enligt detta mått är störst för LäSt ord (maxpoängen är 27 på Bokstavskännedom och 25 på Hur låter orden). I (efter-tests)standardavvikelse räknat är förändringarna också stora, 0,82 för LäSt ord, 1,13 för Bokstavskännedom och 0,70 för Hur låter orden.

För det sistnämnda testet kan vi också jämföra med det normerade resultatet. Hur låter orden genomfördes av 468 elever i förskoleklass i Eskilstuna kommun. Medelpoängen i starten av förskoleklass för denna grupp var 13,8 (standardavvikelse = 6,9) och i slutet av förskoleklass 20,9 (standardavvikelse = 5,1) (Häggström & Brännström, 2008). Ingen av grupperna som får interventionen i förskoleklass gör detta test i starten av förskoleklass, men genomsnittet vid slutet av förskoleklass är 20,8, det vill säga nästan precis samma resultat som för normeringsgruppen. LäSt är inte normerat i förskoleklass, men faktum är att interventionsgruppens genomsnitt 10,8 poäng hamnar i 40-45 percentilen (percentil 1 har lägst resultat, percentil 100 har högst, men resultaten redovisas i intervaller om 5 percentiler) för den normering som gjorts i första klass. Genomsnittsåldern för eleverna vid testtillfället som ligger till grund för normeringen är ungefär 89 månader/7,4 år. I Stanine-poäng motsvarar detta mellan 4 och 5 (Elwér m fl, 2016).¹⁵ Interventionsgruppen presterar därmed redan i förskoleklass nära genomsnittet för elever som är representativa för svenska elever i första klass.

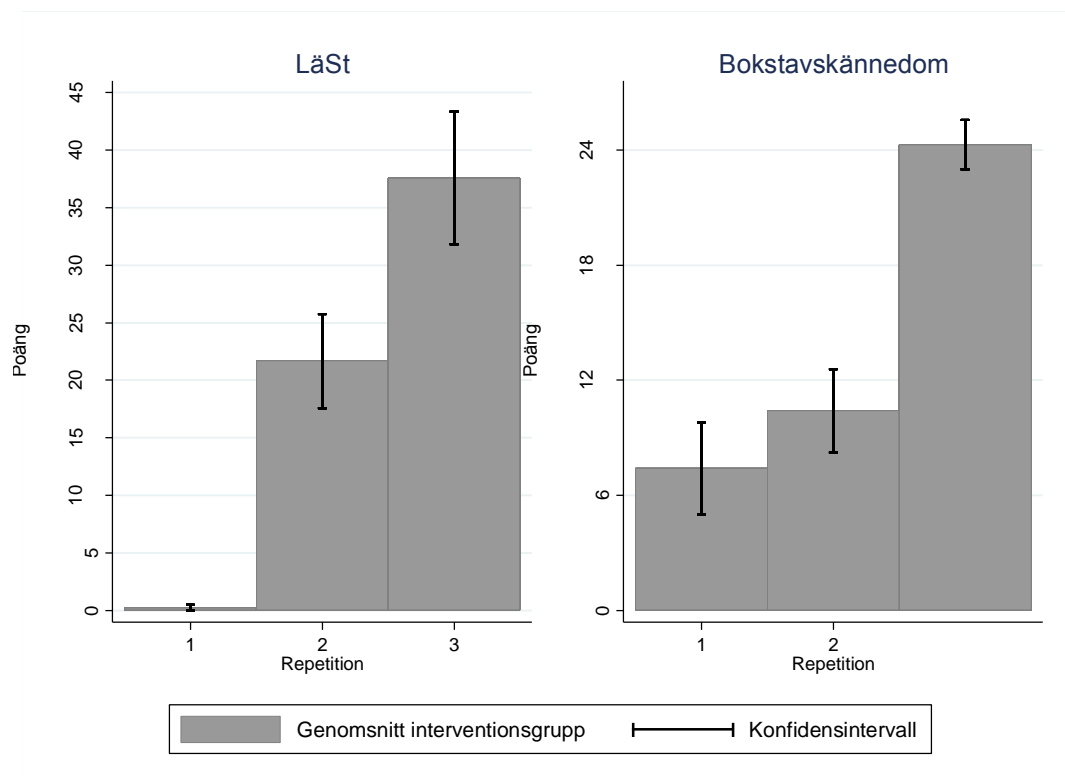
Figur 2 visar utvecklingen för eleverna som fått interventionen i första klass (samtliga fick interventionen individuellt). Det vänstra diagrammet visar utvecklingen över de tre första repetitionerna av LäSt ord. Interventionen sker första terminen i första klass, vilket är mellan första och andra repetitionen. Här sker en förbättring med 21,0 poäng mellan för- och efter-testet. Förbättringen fortsätter också efter interventionens slut, genomsnittet är ytterligare 16,5 poäng högre vid repetition 3. Både ökningen mellan repetition 1 och 2 och mellan 2 och 3 är signifikanta. Det högra diagrammet visar utvecklingen över de tre första repetitionerna av Bokstavskännedom. Här sker de två första repetitionerna i slutet av förskoleklass och början av första klass; interventionen görs därmed mellan repetition 2 och 3. Vi ser igen en relativt stor ökning efter interventionen (12,5 poäng), mycket större än ökningen mellan repetition 1 och 2. Detta visas också av att genomsnittet vid repetition 3 är signifikant högre än vid repetition 2, medan genomsnittet inte är signifikant högre vid repetition 2 än vid repetition 1. Nästan hälften (49 procent) av eleverna når maxpoängen 27 vid repetition 3. Räknat i standardavvikelse är förbättringen från för- till efter-test 1,87 på LäSt ord och 1,93 på Bokstavskännedom. I bägge fallen är därmed förändringarna relativt mycket större än förändringarna mellan för- och efter-test för de elever som fick interventionen i förskoleklass.

För gruppen som får interventionen i första klass kan vi jämföra interventionsgruppens resultat med resultatet från normeringen av LäSt. Repetition 2 är den testomgång som gör att interventionsgruppens ålder ligger närmast genomsnittsåldern för de elever normeringen är gjord på (dessa är som nämnts 89 månader/7,4 år gamla). Medelvärdet på repetition 2, 21,4 poäng, motsvarar

¹⁵ Stanine 4 motsvaras av intervallet 4-10 poäng, stanine 5 motsvaras av intervallet 11-30 poäng.

stanine 5 och hamnar i den 51-55 percentilen (Elwér m fl, 2016). I genomsnitt har därmed eleverna i interventionsgruppen lika bra eller bättre resultat än den genomsnittlige eleven i en grupp som är representativ för svenska elever i första klass.

Figur 2. Utveckling i gruppen som fick interventionen i första klass



Resultat från skattningar med kontrollgrupper

I detta avsnitt redovisar vi resultaten från de analyser där vi använder en kontrollgrupp och DID-modeller för att skatta effekterna av Sofias lässtart enligt beskrivningen i avsnitt Analys.

Vi visar först resultatet från ett så kallat ”balanstest” av hur lika interventions- och kontrollgrupperna är på för-testen. Testet undersöker om vi med hjälp av resultaten på de tre för-testen, individuellt och gemensamt, kan förutsäga vilka elever som får interventionen. Om det går att förutsäga vem som får interventionen är det en indikation på att grupperna inte är lika innan interventionen och att det de därför kanske också skulle ha haft olika utveckling, även om interventionen inte hade inträffat. Vi testar detta i logit modeller där den beroende variabeln är en indikator för om en elev fått interventionen och de tre oberoende variabelerna är resultaten på respektive för-test. Om de beroende variabelerna är statistiskt signifikanta, antingen individuellt eller gemensamt, så har för-testen förmåga att förutsäga vem som får interventionen i högre grad än vad som är önskvärt.

Tabell 3 visar resultatet från dessa modeller. Urvalet i kolumn (1) är alla elever som fått interventionen i grupp, samt de elever från Dalhemsskolans årgångar 2013-2015 som inte fått interventionen varken i förskoleklass eller första klass. I kolumn (2) använder vi samma urval, men

viktar med vikterna från CEM. Kolumn (3) och (4) visar resultat för interventionsgruppen som fått Sofias lässtart individuellt och dess kontrollgrupp. Kolumn (4) tar då bara med de elever som har noll poäng på LäSt ord för-testet, vilket samtliga i kontrollgruppen har. LäSt kan därför inte ingå i dessa två modeller.

Tabell 3. *Balanstest av interventions- och kontrollgrupper*

Variabler	(1) Grupp	(2) Grupp + match	(3) Individuell	(4) Individuell + 0
<i>LäSt</i>	-0,092*** (0,032)	-0.026 (0,059)		
<i>Bokstavskännedom</i>	0,038 (0,35)	0,023 (0,038)	-0,026 (0,070)	-0,081 (0,083)
<i>Hur låter orden</i>	0,11 (0,079)	0,11 (0,10)	0,034 (0,052)	0,007 (0,059)
<i>Konstant</i>	-3,25** (1,45)	3,11 (2,17)	0,073 (0,69)	0,33 (0,74)
Totalt antal elever	108	86	32	30
<i>p</i> -värde, gemensam sign.	0,04	0,65	0,79	0,62
Pseudo- <i>R</i> ²	0,06	0,02	0,01	0,02

Not: Robusta standardfel i parentes, $p < 0.01$ ***, $p < 0.05$ **, $p < 0.1$ *. I kolumn (1) och (2) består interventionsgruppen av elever från årgång 2015 som fått interventionen i grupp. I kolumn (1) är kontrollgruppen samtliga elever i årgångarna 2013-2015 på Dalhemsskolan som inte fått interventionen. I kolumn (2) använder vi CEM för att matcha fram en kontrollgrupp. Observationerna i skattningen i kolumn (2) är också viktade med vikter från CEM. I kolumn (3) och (4) utgörs interventionsgruppen av elever i årgång 2015 som fått interventionen individuellt i förskoleklass. Kontrollgruppen i kolumn (3) är samtliga elever i årgång 2013-2014 som fått interventionen individuellt i första klass. I kolumn (4) begränsar vi både interventions- och kontrollgrupp till de som hade noll poäng på LäSt ord för-testet. *p*-värdet i nedre delen av tabellen kommer från ett Wald-test av hypotesen att variablerna är gemensamt signifikanta.

Det är endast i första kolumnen som vi finner någon statistisk signifikans, i de andra skattningarna är koefficienterna inte signifikant skilda från noll, varken individuellt eller gemensamt. Detta pekar på att urvalet i kolumn (2) har större chanser att fånga de kausala effekterna av gruppinterventionen och att vår matchningsprocedur ökar balansen mellan interventions- och kontrollgruppen.¹⁶ Eftersom alla i kontrollgruppen i skattningarna av den individuella interventionens effekter har noll poäng på LäSt, så har vi vidare större förtroende för att urvalet i kolumn (4) fångar effekterna av interventionen.

Tabell 4 innehåller våra huvudresultat och visar koefficienter, standardfel (i parentes) och Hedges' *g* (i hakparentes) från skattningar med urval som motsvarar kolumn (1) till (4) i Tabell 3. Urvalsstorlekarna skiljer sig något åt, eftersom enstaka elever inte gjort samtliga tre förtest. Panel A innehåller resultaten för LäSt, Panel B resultaten för Bokstavskännedom och Panel C visar resultaten för Hur låter orden.

¹⁶ Notera att nivån på för-testet strikt sett inte måste vara lika för att våra DID-modeller ska skatta de kausala effekterna av interventionen, antagandet som måste gälla säger att trenden skulle varit densamma om interventionsgruppen inte skulle fått interventionen (se tidigare avsnittet *Analys*). I vårt fall är det dock troligt att nivån innan interventionen är en viktig förklaring också till utvecklingen över tid.

Tabell 4. Interventionseffekter i förskoleklass från difference-in-differences modeller

Panel A: LäSt ord				
Variabler	(1) Grupp	(2) Grupp + match	(3) Individuell	(4) Individuell + 0
<i>Intervention</i>	2,39 (1,64) [0,14]	3,29* (1,66) [0,31]	2,63*** (0,747) [0,91]	1,94*** (0,470) [1,24]
<i>Konstant</i>	8,32*** (1,08)	7,68*** (1,09)	0,059 (0,059)	0,059 (0,059)
Totalt antal elever	109	86	33	31
Interventionsgrupp	31	30	16	14
R^2	0,015	0,05	0,30	0,41
Panel B: Bokstavskänedom				
Variabler	(1) Grupp	(2) Grupp + match	(3) Individuell	(4) Individuell + 0
<i>Intervention</i>	2,02** (0,976) [0,33]	1,11 (1,08) [0,21]	2,88** (1,22) [0,42]	2,85** (1,13) [0,44]
<i>Konstant</i>	5,14*** (0,564)	6,23*** (0,721)	4,87*** (0,949)	4,87*** (0,951)
Totalt antal elever	108	86	31	29
Interventionsgrupp	31	30	16	14
R^2	0,04	0,01	0,16	0,19
Panel C: Hur låter orden				
Variabler	(1) Grupp	(2) Grupp + match	(3) Individuell	(4) Individuell + 0
<i>Intervention</i>	0,342 (0,700) [0,09]	0,339 (0,705) [0,12]	-0,050 (1,56) [0,009]	0,343 (1,70) [0,07]
<i>Konstant</i>	2,08*** (0,405)	1,89*** (0,434)	5,80*** (0,840)	5,80*** (0,842)
Totalt antal elever	109	86	31	29
Interventionsgrupp	31	30	16	14
R^2	0,00	0,00	0,00	0,00

Not: Robusta standardfel i parentes, $p < 0.01$ ***, $p < 0.05$ **, $p < 0.1$ *. Hedges' g i hakparenteser under estimaten av interventionseffekten. I kolumn (1) och (2) består interventionsgruppen av elever från årgång 2015 som fått interventionen i grupp. I kolumn (1) är kontrollgruppen samtliga elever i årgångarna 2013-2015 på Dalhemsskolan som inte fått interventionen. I kolumn (2) använder vi CEM för att matcha fram en kontrollgrupp. Observationerna i skattningen i kolumn (2) är också viktade med vikter från CEM. I kolumn (3) och (4) utgörs interventionsgruppen av elever i årgång 2015 som fått interventionen individuellt i förskoleklass. Kontrollgruppen i kolumn (3) är samtliga elever i årgång 2013-2014 som fått interventionen individuellt i första klass. I kolumn (4) begränsar vi både interventions- och kontrollgrupp till de som hade noll poäng på LäSt ord för-testet.

Effekterna av gruppinterventionen är positiva på alla tre test och i båda specifikationerna i kolumn (1) och (2). Koefficienten för interventionsgruppen är dock bara signifikant (på 10-procentsnivån) i kolumn (2), Panel A (för LäSt ord), och kolumn (1), Panel B (för Bokstavskännedom). Effektstorlekarna är samtidigt långtifrån obetydliga: de är 0,14 och 0,32 för LäSt ord och 0,33 och 0,21 för Bokstavskännedom. Med vår urvalsstorlek har vi förmodligen inte tillräcklig statistisk styrka för att hitta effekter av den storleksordningen med rimlig säkerhet. Både koefficienter och effektstorlekar är mindre när vi använder Hur låter orden. Ingen koefficient är signifikant skild från noll och effektstorlekarna är 0,09 och 0,12.

Skattningarna av interventionseffekten för eleverna som fått Sofias lässtart individuellt i förskoleklass är positiva, stora och signifikanta mätt med LäSt ord och Bokstavskännedom i kolumn (3) (Panel A och B). När vi bara inkluderar eleverna med noll poäng på LäSt ord i kolumn (4) ökar effektstorleken för båda testen. Effektstorlekarna är 0,91 och 1,24 på LäSt ord och 0,42 och 0,44 på Bokstavskännedom. Båda dessa skillnader är fortfarande signifikanta om vi istället använder ett icke-parametriskt Wilcoxon-test, som testar om fördelningarna av förändringen från för- till efter-test skiljer sig åt mellan interventions- och kontrollgrupp.¹⁷ Däremot är effekten nära noll för båda grupperna på Hur låter orden (se Panel C).

I jämförelse är därför effektstorlekarna för åtminstone LäSt och Bokstavskännedom betydligt större för de elever som fått interventionen individuellt än för de som fått den i grupp. Detta beror inte på skillnader i antalet sessioner de båda grupperna har fått, eleverna som fått gruppinterventionen har i genomsnitt fått fler sessioner. Ingen av grupperna har i genomsnitt fått de 30 sessioner som är det avsedda antalet och i interventionsgruppen är det genomsnittliga antalet sessioner bara 15,6. Däremot finns det stora skillnader i resultaten på för-testen mellan grupperna. Eleverna som fått interventionen individuellt har generellt mycket lägre resultat, men det finns också elever i gruppinterventionen som har låga resultat innan interventionen. För att undersöka om det är denna skillnad mellan grupperna som avspeglas i de skilda resultaten visar vi i Figur 3 förändringen från för- till efter-test för elever i bägge grupperna som har noll poäng på LäSt ord och färre än åtta poäng på Bokstavskännedom på för-testet. Vi har valt dessa gränser efter diskussioner med specialpedagogerna som utfört interventionen, som menar att elever med sådana resultat har en förhöjd risk att behöva specialpedagogiska insatser efter förskoleklass. Detta ger oss 12 elever som fått en individuell intervention och 6 elever som fått en gruppintervention. Det bör påpekas direkt att medan genomsnittet på för-testet i grupperna är detsamma per definition på LäSt ord, så har gruppinterventionseleverna ett klart högre genomsnitt på för-testen i både Bokstavskännedom och Hur låter orden. Grupperna riskerar därmed fortfarande att inte vara helt jämförbara.

Figur 3 indikerar att eleverna som fått gruppintervention har gjort klart större förbättringar på LäSt ord. Genomsnittsförändringen (och genomsnittet eftersom alla har 0 poäng på första testet) för gruppinterventionseleverna är 8,5 medan det är 1,8 för eleverna med individuell intervention. Skillnaden mellan fördelningarna är signifikant i ett icke-parametriskt Wilcoxon-test (p -värde =

¹⁷ Fördelen med detta test jämfört med våra tidigare skattningar är att det inte gör något antagande om formen på sambandet mellan interventionen och testresultat (sambandet behöver inte vara linjärt) och hur standardfelen är fördelade (de behöver inte vara normalfördelade). Nackdelen med testet är att vi inte får något estimat av storleken på effekten.

0,004). Även på Bokstavskännedom har gruppinterventionerna förbättrat sig mer. Respektive genomsnittsförändring är 12,7 (grupp) och 7,5 (individuell) och även denna skillnad är signifikant i ett Wilcoxon-test (p -värde = 0,004). Förhållandet mellan grupperna är sedan det omvända på Hur låter orden. Gruppen med individuell intervention förbättrar sig i genomsnitt med 5,5 poäng medan gruppinterventionseleverna förbättrar sig med i genomsnitt 1,2 poäng. Även denna skillnad är signifikant i ett Wilcoxon-test (p -värde = 0,06). Gruppinterventionseleverna ligger dock så pass högt på för-testet (genomsnittet är 21,7 och maxpoängen är 25) att det finns risk för takeffekter.

Vi diskuterar fler möjliga anledningar till skillnaderna i resultat, samt för en mer generell diskussion om hur resultaten kan tolkas i nästa avsnitt.

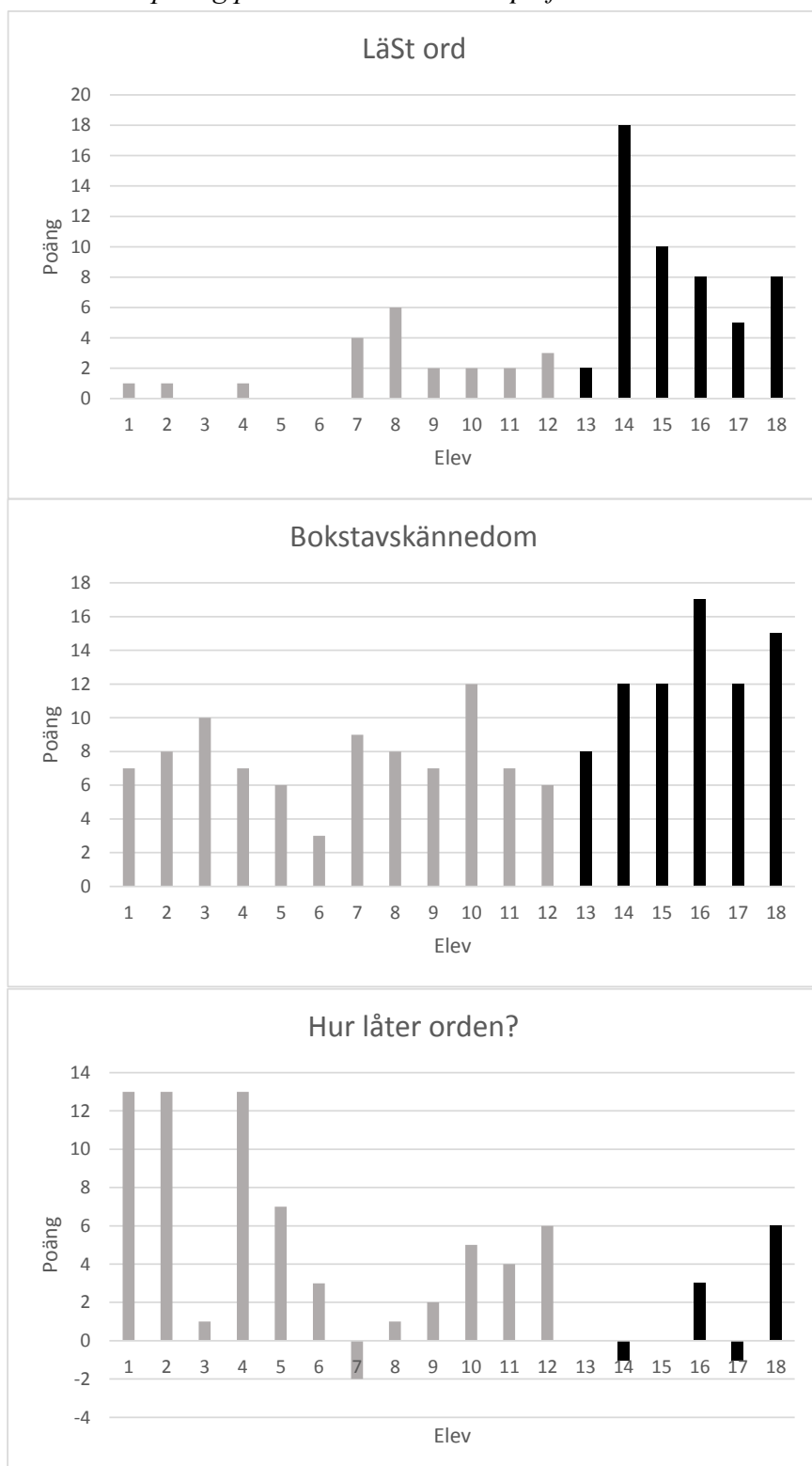
Diskussion

Våra resultat indikerar att elever som har fått Sofias lässtart har bättre resultat än elever i en jämförbar kontrollgrupp, som inte fått interventionen. För att sätta effekterna av interventionen i perspektiv skulle vi vilja jämföra med liknande svenska interventioner i förskole- och första klass. Vi känner dock inte till någon tidigare utvärdering av en intervention riktad mot en liknande målgrupp som de elever som fått Sofias lässtart. Så även om man bör vara försiktig med att jämföra effekter från interventioner från andra länder och/eller som beräknats med olika typer av test som grund, så är det i detta fall den bästa tillgängliga jämförelsen. För de elever som fått interventionen individuellt i förskoleklass är effektstorlekarna på LäSt och Bokstavskännedom större, eller mycket större, än de genomsnittliga effekter som hittats för liknande interventioner i synteser av den internationella forskningslitteraturen, som undersöker interventioner för elever i liknande målgrupper och där effektstorlekarna beräknats med liknande typer av test (se Slavin m fl, 2011; Dietrichson m fl, 2015).¹⁸ I båda dessa översikter har tutoring-interventioner de största effektstorlekarna (i genomsnitt omkring 0,3-0,4). Effekterna av Sofias lässtart är också ungefär lika stora för Bokstavskännedom och mycket större för LäSt ord, som effekterna i Wolff (2011), där målgruppen är svenska tredjeklasselever med lässvårigheter.

Effekterna är mindre för eleverna som får Sofias lässtart i grupp på både LäSt ord och Bokstavskännedom än för eleverna som fått interventionen individuellt och i många fall är skillnaden mot kontrollgrupperna inte statistiskt signifikanta. Effektstorlekarna är dock fortfarande relativt stora i förhållande till genomsnitten i den internationella litteraturen och i paritet med de som presenteras i Wolff (2011). Det är därför möjligt, som vi diskuterar nedan, att urvalsstorleken är för liten för att vi ska finna signifikanta resultat.

¹⁸ Denna litteratur domineras kraftigt av interventioner genomförda i USA.

Figur 3. Förändring mellan för- och efter-test för interventionselever som hade 0 poäng på LäSt ord och ≤ 8 poäng på Bokstavskänedom på för-testet.



Not: De gråa staplarna representerar förändringen mellan för- och efter-test för elever som fått interventionen individuellt i förskoleklass, medan de svarta staplarna visar samma förändring för eleverna som fått gruppinterventionen i förskoleklass. Varje stapel representerar en enskild elevs förändring.

Vidare är elevgrupperna som får interventionen i grupp och individuellt troligen inte jämförbara. I regel har de elever med störst svårigheter fått interventionen individuellt. Om Sofias lässtart har störst effekter för elever med de största svårigheterna så är det möjligt att vi hade sett liknande resultat om dessa fått en gruppintervention. Att gruppinterventionseleverna gör större förbättringar när vi undersöker den grupp av elever som har allra lägst resultat innan interventionen tyder på att så kan vara fallet. Eftersom resurserna som krävs per elev för att genomföra interventionen minskar ganska dramatiskt med gruppstorleken skulle det vara en intressant hypotes att testa i framtida studier.

Ingen av interventionsgrupperna visar upp några signifikanta framsteg gentemot kontrollgrupperna på testet Hur låter orden. En möjlig anledning till detta kan vara att de områden inom läsning som testas med Hur låter orden utgör en betydligt större del av den vanliga undervisningen i förskoleklass på de skolor som ingår i undersökningen. Eftersom Bornholmsmodellen inriktar sig på samtliga delar av Hur låter orden medan Sofias lässtart fokuserar delar av innehållet i testet så är det möjligt att kontrollgruppen fått mer undervisning i de andra två delarna. Åtminstone vissa elever interventionsgruppen har också relativt höga resultat redan på för-testet, så takeffekter är också en möjlig delförklaring.

Vi jämförde också utvecklingen inom interventionsgrupperna före och efter interventionen och jämförde i sin tur denna utveckling med normerade resultat. Denna analys inkluderade också de elever som fått interventionen i första klass (för vilka vi annars inte har någon kontrollgrupp). Båda grupperna uppvisar stora förbättringar över tid och har resultat i paritet med genomsnittet hos eleverna som normeringen gjorts på. Med tanke på att skolorna i vårt urval är långt ifrån normalskolor och gruppen som får interventionen är de barn med störst svårigheter i dessa skolor, är detta ett anmärkningsvärt starkt resultat.

Då eleverna som fått Sofias lässtart i första klass ser ut förbättra sig mer i för-efter analysen hade det varit mycket intressant att jämföra dessa elever med en kontrollgrupp, vilket vi tyvärr inte hade möjlighet att göra. En möjlig förklaring till skillnaden skulle kunna vara att elever i allmänhet upplever en kraftigare progression i första klass än i förskoleklass. Men två andra saker skiljer sig också mellan de som fått interventionen i förskole- och första klass: 1) alla som fått interventionen i första klass fick den individuellt; 2) förstaklasseleverna har i genomsnitt fått ca 10 sessioner mer. Förskoleklass-eleverna har i genomsnitt endast fått lite mer än hälften (15,6) av det tänkta antalet sessioner (som är 30) när efter-testen genomförs. Dessa förklaringar skulle kunna bidra till att utvecklingen ser bättre ut i första klass, men avsaknaden av kontrollgrupp gör det svårt att uttala sig om orsakerna med någon större säkerhet. En anledning till att förskoleklass-eleverna inte fått mer än ungefär hälften av sessionerna är att interventionen kom igång relativt sent på Jens Billeskolan och Raus Planterings skola; eleverna där gjorde inte för-testen förrän i april respektive i maj.

Vår analys har även mer generella begränsningar, som i stor utsträckning bottnar i att urvalsstorleken är liten. En liten urvalsstorlek ger upphov till åtminstone två problem. För det första så har vi en begränsad statistisk styrka till att hitta signifikanta effekter i den matchade DID designen. Detta kan vara anledningen till att resultaten för gruppinterventionen som regel inte är signifikanta. Effektstorlekarna är som nämnts långtifrån obetydliga för denna grupp. Det andra

problemet är att så kallade falska positiva effekter också är mer sannolika (se t ex Christley, 2010). Det vill säga, risken att de stora effekter vi ser för gruppen som fått interventionen individuellt är av slumpmässig art istället för sanna effekter ökar, ju mindre urvalet av elever är.

Att antalet skolor, årgångar och klasser är litet skapar också risker att effekterna av interventionen blandas samman med andra faktorer som påverkar elevernas resultat. Hade vi kunnat välja matchade kontrollgrupper från ett större antal skolor, årgångar och klasser, eller ännu bättre kunnat randomisera elever till interventions- och kontrollgrupper, hade risken varit mindre att faktorer som skolkvalitet, elevunderlag och lärarkvalitet snedvrider resultaten. Att vi kan skapa kontrollgrupper som inte skiljer sig från varandra på för-testen, som alla dessa faktorer har påverkat, torde dock minska dessa risker och det är oklart om skillnader i dessa faktorer gör att vi överskattar eller underskattar storleken på effekterna. Eftersom svenska elever inte gör några gemensamma test förrän de nationella proven i tredje klass är det svårt att i efterhand skapa breda kontrollgrupper för yngre åldersgrupper.

Ett ytterligare problem rör statistisk signifikans och bottnar i att eleverna går tillsammans i skolor och klasser. Standardfelen för den genomsnittliga effekten riskerar därför att underskattas eftersom elevernas resultat inte nödvändigtvis behöver vara oberoende av varandra. Eftersom antalet skolor/klasser ("kluster") är så pass litet är det dock inte enkelt att hitta en fungerande metod för att beräkna "kluster-robusta" standardfel. Som vi nämner ovan gör samtidigt den lilla urvalsstorleken att vi har svårt att finna även relativt stora effekter med rimlig statistisk styrka och den bästa lösningen på alla dessa problem är därför att genomföra en större studie av Sofias lässtart.

Slutsatser

Trots de begränsningar vi nämner ovan bör resultaten vi har redovisat i denna rapport betraktas som mycket lovande. Sofias lässtart verkar kunna öka läsförmågan hos elevgrupper som i regel har det svårt med läsningen och som riskerar att hamna efter sina kamrater redan tidigt i skolgången. Även om interventionen, särskilt i individuell form, är relativt intensiv så krävs mindre resurser än i många liknande interventioner. Detta framförallt på grund av att interventionen genomförs under kort tid, ca 8 veckor, vilket är mindre än hälften så långt som medel- och medianlängden för liknande tutoring-interventioner i Dietrichson m fl (2016). Eftersom de skattade effekterna av interventionen är i paritet med de interventioner som har allra störst effekt, och större än genomsnittet av tutoring-interventioner (som i sin tur är den interventionstyp som har allra störst effekter), skulle Sofias lässtart även kunna vara ett mycket resurseffektivt sätt att hjälpa elever med lässvårigheter.

Referenser

- Abadie, A., 2005. Semiparametric difference-in-differences estimators. *Review of Economic Studies*, 72(1), 1–19.
- Björklund, A., Fredriksson, P., Gustafsson, J-E., & Öckert, B. (2010). *Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?* IFAU Rapport 2010:3.
- Björklund, A., & Salvanes, K. (2011). Education and family background. In Hanushek, E. A., Machin, S., & Woessmann, L. (eds.), *Handbook of the Economics of Education, Volume 3*, 201-247.
- Christley, R. M. (2010). Power and error: Increased risk of false positive results in underpowered studies. *The Open Epidemiology Journal*, 3, 16-19.
- 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.
- Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A-M. (2016). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. Unpublished manuscript.
- Elwér, Å., Fridolfsson, I., Samuelsson, S., & Wiklund, C. (2016). *LäSt – Test i läsförståelse, läsning och stavning för åk 1-6*. Hogrefe Psykologiförlaget: Stockholm.
- Heckman, J.J. (2000). Policies to foster human capital. *Research in Economics*, 54(1), 3-56.
- Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5), 1017-1098.
- 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.
- Häggström, I. (2007). Att förebygga läs- och skrivsvårigheter med språklekar. I Granström, K. (red) *Forskning om lärares arbete i klassrummet, Forskning i fokus, nr 33*, Myndigheten för Skolutveckling.
- Häggström, I, & Brännström, L. (2008). *Hur låter orden*. Ing-Read AB: Linköping.
- Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106(493), 345-361.
- Johansson, M-G. (2009). *LäsEttan*. Natur & Kultur: Stockholm.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied Social Research Methods Series, v. 49.
- Lipsey, M. W., Puzio, K., Yun, C., Herbert, 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*. Institute of Education Sciences Report, NCSER 2013-3000.
- 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.
- OECD (2010). *PISA 2009 results: Overcoming social background – Equity in learning opportunities and outcomes (Volume II)*. PISA, OECD Publishing.
- OECD (2013). *PISA 2012 results: Excellence through equity – Giving every student the chance to succeed (Volume II)*. PISA, OECD Publishing.
- 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.
- Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417-453.
- Scammaca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4-12: 1980-2011. *Journal of Learning Disabilities*, 48, 369-390.
- SCB (2016). www.statistikdatabasen.scb.se. Data hämtade 2016-07-10.
- Skolverket (2016). www.skolverket.se/statistik-och-utvardering/statistik-i-databaser. Data hämtade 2016-07-10.
- Slavin, R. E., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6, 1-26.
- Slavin, R. E., & Madden, N. (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness*, 4, 370-380.
- Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer*. Skolverket: Stockholm.
- 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*. The National Academies Press: Washington DC.
- SOU (2012). *Läsandets kultur – Slutbetänkande av Litteraturutredningen*. Statens offentliga utredningar, SOU 2012:65.
- Snowling, M., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, 81, 1-23.
- Wanzek, J., Vaughn, S., Scammaca, N. K., Metz, K., Murray, C. S, Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, 83, 163-195.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91, 461-481.
- Wolff, U. (2011). Effects of a randomised reading intervention study: An application of structural equation modelling. *Dyslexia*, 17, 295-311.
- Wolff, U. (2016). Effects of a randomized reading intervention study aimed at 9-year-olds: A 5-year follow-up. *Dyslexia*, 22, 85-100.